THE UNIVERSITY *of* EDINBURGH

The collective dynamics of self-propelled particles in confining environments

Elliot J Marsden

Doctor of Philosophy
The University of Edinburgh
2016

Abstract

Self-propelled particles are a class of far-from-equilibrium systems which present many complex, emergent features that are not obvious from the microscopic dynamics. Simulations of well-chosen instances of such systems are a powerful yet tractable method of investigating many real-world phenomena. The frequently non-time-reversible interactions of many cases of self-propelled particles with surfaces means that the environment has an impact on large-scale behaviour in a way that would not be true for particles close to thermal equilibrium.

This work investigates several examples of such systems, and compares them with experimental results for comparable systems: firstly, the spatial distribution of smooth-swimming mutants of *Eschericia Coli* within water-in-oil emulsion is investigated, and its dependence on inter-bacterial interactions and the size of water droplets. The nature of bacterial collisions is inferred through data analysis and simulation. Secondly, pattern formation by chemotactic run-and-tumble bacteria due to secretion of a chemoattractant by the bacteria themselves, demonstrating a range of approaches to control the formation of biofilms by bacteria. Finally the dependence of the bulk transport properties of chemotactic self-propelled particles in porous environments, on their detailed dynamics, is probed: how they interact with obstacles, their form of chemotactic response, their ability to actively enhance their rotational noise, and their method of sensing chemical gradients.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

Parts of this work have been published in [1].

Parts of this work have been published in [2].

(*Elliot J Marsden, 2016*)

Acknowledgements

Many thanks to Mike Cates and Davide Marenduzzo for supervising this work and for consistent source of motivation throughout. I am also thankful to my collaborator Dana Vladescu for her valuable experimental expertise and feedback. Thanks also to Wilson Poon for valuable advice and criticism, and to Chantal Valeriani for enthusiastic feedback.

I also greatly appreciate useful discussions with my colleagues, in particular Tim Bush, Fergus Cullen, Elizabeth Eardley, Catherine Mills, Toby Searle and Joe Forth.

The work was funded by the UK Engineering and Physical Sciences Research Council.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In Chapter 2, we explain the biological background motivating the work that follows. We explain the strategies small organisms use to be able to swim in the viscous world in which they live. We proceed to outline the importance of surfaces to the survival of bacteria in harsh conditions, and describe the physical consequences of swimming on their interaction with these surfaces. We review experiments exploiting this interaction to show the effect that a static environment can have on steady-state distributions of bacteria, and show the reason for this deviation from the predictions of equilibrium statistical physics.

In Chapter 3, we describe the theoretical techniques we will use throughout the thesis. We give a brief overview of Brownian motion in terms of its causes and implications, and extend this to the case of rotational Brownian motion: a dominant effect in the bulk motion of self-propelled particles for the parameters of relevance here. We discuss the importance of the length scale of swimmers on their ability to swim, and on the useful simplifications which a low-Reynolds number context allows us to make. We explain numerical techniques used to solve partial differential equations, and the method of operator splitting, leading to an outline of efficient methods used to solve ordinary differential equations. Finally, we describe the Metropolis-Hastings algorithm, and show how it may be modified to generate high-density packings of geometric objects in bounded periodic or non-periodic spaces.

In Chapter 4, we analyse the behaviour of motile bacteria within a sphere. We describe the experimental arrangement used to realise this system, and justify the methods used to analyse the experimental data so obtained. We then proceed

to motivate an analytic model of the system, and fit the data to this model to determine the strength of inter-bacterial scattering induced by a collision. This is followed by the description of our simulation model, and the results it predicts compared with experiment. Fitting to the experimental data is again performed, resulting in a microscopic parameter describing inter-bacterial scattering. The spatial data provided by the agent-based model allows for further comparison in terms of statistical moments and pair-wise correlations, which serves to confirm the earlier results.

In Chapter 5 we study the distribution of autochemotactic bacteria in confining spaces. We first outline our model of the bacteria and chemoattractant, then introduce and justify the measure we use to quantify the homogeneity of bacterial distributions. This is followed by a study of the model in one dimension, demonstrating and quantifying the existence of a point of instability of a uniform distribution, and a mapping onto a coarse-grained analytic set of equations. We proceed to extend the model to two dimensions, showing how the same point of instability is modified in higher dimensions. We modify the environment to contain walls in various configurations, showing how the point and form of clustering transition is modified in such circumstances.

In Chapter 6, we investigate the effect on each qualitative choice of swimming dynamics on diffusion and migration in porous media. We begin by outlining the biological mechanism of chemotaxis in *E. coli* and similar organisms, and transform this into a mathematical model of time-averaged chemotaxis using a 'chemical memory'. This is then calibrated with the non-time-averaged analogue to make the two comparable. We show the results of simulations to measure migration efficiency in a linear chemical gradient, before giving a description of the model of porous media that was implemented. The effect of the porosity of this environment on drift and diffusion are shown, and the impact of time-averaging is demonstrated through calculating deviations from the fluctuation-dissipation theorem.

Finally, in Chapter 7, we summarise the work as a whole.

Chapter 2

Background

In the following chapter we introduce the background biology relevant to the work contained in subsequent chapters. We also review relevant experimental and theoretical work carried out in the last few years which has motivated much of the work shown later.

2.1 Directed motion in small organisms

A large class of organisms control their position in the environment by acting to achieve a non-zero velocity over long periods of time. This helps to ensure a constant source of nutrients in static environments such as quiescent fluids.

The methods used to achieve this motion, known as 'motility', vary between organisms. At small scales, many intuitively effective mechanisms fail due to the dominance of viscous forces. The classic example is that of a scallop, which moves by opening then rapidly shutting its shell, so as to produce a jet of water whose momentum propels the scallop forwards. At small scales, the opening and shutting actions 'undo' the effect of each other, so as to provide no net displacement. In general, sufficiently small organisms in fluids cannot swim using any method of deforming their bodies that is time-reversible: this is in fact known as the 'scallop theorem' [3].

Of the options that remain for small swimmers, we outline two important classes of motion that are used in nature. The first uses a large number (on the order

of 10^3 [4]) of long, thin hair-like structures known as 'cilia', which waft in an undulating manner similar to the breast-stroke (a time-irreversible deformation). This tactic is typically adopted by larger organisms, specifically the class of eukaryotes named for this feature, 'ciliates'. These range from 10 µm to 4000 µm in size [5], and can achieve speeds on the order of $\mathrm{mm\,s^{-1}}$.

A second strategy also involves long external appendages, but a much smaller number: in the range 1 to 10. Instead of an undulatory action, these rotate in a propellor-like way (this is also a time-irreversible process since the direction of rotation would reverse upon time reversal). These structures are known as 'flagella', and are typically found in bacteria, of sizes 1 µm to 20 µm. They typically achieve speeds on the order of $10\,\mathrm{\mu m\,s^{-1}}$. The flagella themselves are rigid helices, 20 nm thick [6], with a motor at their base that can rotate in either direction, in a corkscrew-like motion.

Within the set of flagellated bacteria, several configurations of the flagella themselves are found. The simplest arrangement is a single flagellum at a single location, which is known as a 'monotrichous' arrangement. This is seen in *Vibrio cholerae*, the species responsible for cholera [7]. If several flagella are found at the same site, this is known as a 'lophotrichous' arrangement, which is seen in *Spirillum volutans* [8]. Species which have flagella at opposite ends of their body are known as 'amphitrichous', such as *Rhodospirillum rubrum*. Finally, species with a uniform distribution of flagella on their surface are known as 'peritrichous', and include large classes such as the *Salmonella* genus [9] and the model organism *Escherichia coli* [10].

These varying distributions of flagella are accompanied by a variety of styles of locomotion. A simple method is used by monotrichous bacteria such as *Pseudoalteromonas haloplanktis* [11], whose single flagellum periodically reverses its direction of rotation. *S. volutans* adopts the same strategy, coordinating this reversal across its many flagella. We refer to this as 'run-reverse' motion.

A more complex strategy is adopted by the marine bacterium *Vibrio alginolyticus* [12], which has a three-step mechanism:

1. Run: The flagellum rotates in one direction to achieve forward motion
2. Reverse: The flagellum reverses direction to achieve backwards motion
3. Flick: The flagellum performs a 'flicking' motion which randomises its

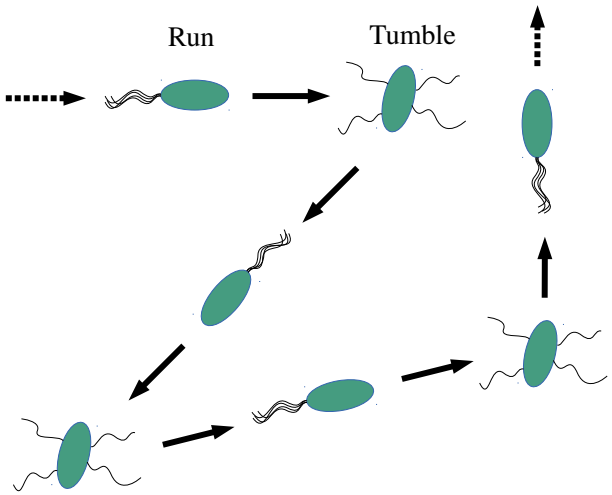


Figure 2.1 *Schematic of run-and-tumble dynamics, with its mechanism shown for a flagellated swimmer. Straight runs aided by coherent, bundled flagella are interrupted by short periods of random rotation caused by the incoherent, independent rotation of the separate flagella.*

direction

The mechanism typically adopted by peritrichous bacteria is so-called 'run-and-tumble' motion. This is similar to the 'run-reverse-flick' mechanism, but is a two-step process, and involves the coordination across the flagella:

1. Run: The flagella rotate in a particular direction, the flagella bundle together and achieve forward motion

2. Tumble: At least some of the flagella reverse direction, the flagellar bundle de-couples, the flagella rotate independently, and the cell body rotates incoherently

This is shown in Fig. 2.1. As mentioned, many important bacteria use this mechanism to explore their environment, such as the model organism *E. coli*, *S. typhimurium* and *Bacillus subtilis*. It is this form of motility on which we will focus henceforth.

2.2 Bacteria living on surfaces

Surfaces are a valuable resource that bacteria may exploit. There are three main reasons for this: shelter, stability, and community. Surfaces provide physical shelter from external dangers such as antibiotic chemicals [13]. Protection from chemicals near surfaces is possible if the route from the source of the chemical to the bacteria is tortuous, for example if antibiotics must percolate through packed, saturated soil from their origin [14, 15]. They also provide a point of anchor when in fast-flowing fluids [16], allowing bacteria to take control of their position in the environment. If a particular region of the environment is both nutrient-rich, and has a local fluid flow, the only options available to bacteria are to abandon the nutrients; perform energetically costly active swimming against the flow; or to form a physical attachment to a nearby surface. An additional advantage of this scenario when compared to localising in a static nutrient-rich environment, is the possibility of a continuous supply of nutrients from upstream of the flow, in environments such as water treatment pipes [17, 18]. A third important advantage of surfaces is that they provide a common meeting point from which to build communities of bacteria that can aid each others' chance of survival [19]. Even in the absence of an explicit advantage of a location, if it is more likely to be visited by bacteria, this increases the possibility of collective behaviour. In this sense, surfaces can break the symmetry of the environment and provide nucleation points for the occurrence of events that require high bacterial number densities [20, 21].

The communal structure that is formed when bacteria come together on a surface is referred to as a 'biofilm' [22, 23] (Fig. 2.2). This is a structure consisting of two components. The first is the bacteria themselves, which are often differentiated in their gene expression profiles from their free-space, so-called 'planktonic' counterparts [24, 25]. For example, 38 % of the genes in *E. coli* were shown to be affected by entering a biofilm state, involving changes such as reducing the production of proteins involved in constructing flagella [26]. The second major component of a biofilm is a matrix made of 'extracellular polymeric substance' (EPS), which is secreted by the bacteria on formation of the biofilm, and provides the 'house of the biofilm cells' [27]. It may be this substance that makes diffusion of antibiotics into biofilms less effective [28].

The importance of these biofilms is in their effectiveness at sustaining populations of bacteria. In some cases these bacteria are welcome. Bioremediation is the process of using organisms to remove unwanted substances from the environment, such as oil spills [29]. In cases of especially recalcitrant targets, biofilms are beneficial due to their ability to increase bacterial populations quickly, and to

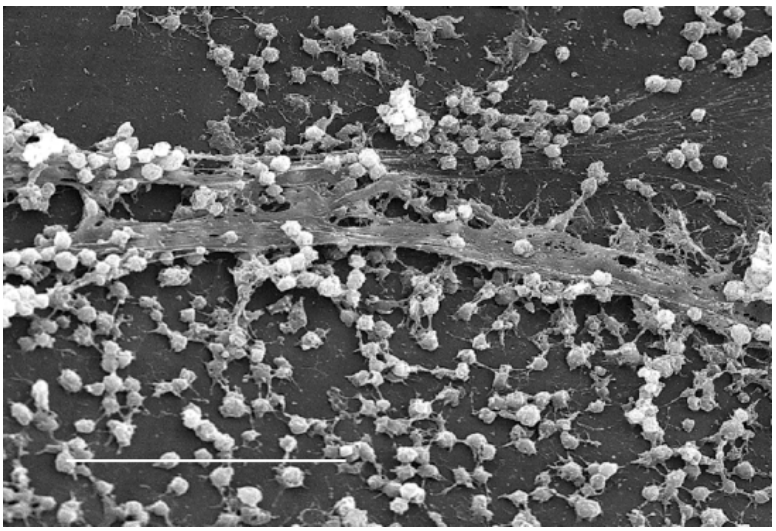


Figure 2.2 *A scanning electron micrograph of a biofilm of Staphylococcus aureus on the surface of a catheter. The image is approximately* $60\,\mu m \times 40\,\mu m$. *Obtained from the CDC Public Health Image Library.*

increase horizontal gene transfer that may improve the rate of degradation of the unwanted substance [30]. In many cases, the formation of a biofilm is a regrettable event. One of the most common causes of the failure of medical implants and prosthetic devices is infection by bacteria [31]. The nano-scale structure of these materials has a large impact on the formation of biofilms: changing its shape can result in a factor of three change in the number of bacteria on the surface [32]. Dental plaques have also been shown to function similarly to biofilms [33].

The events that lead from a uniform distribution of planktonic bacteria to a dense surface community depend on the bacterial species in question [34], however in all cases an initial increase in density is required, to the extent that a threshold cluster of bacteria can directly sense each other. In many cases this is the result of random fluctuations in the bacteria's positions, however there are processes that can make this fluctuation more likely to occur. The threshold density at which bacteria can sense each other may be lowered by their advertising of their presence through the secretion of extracellular molecules, which then diffuse and are detected by others nearby. This coupling of the reception of external, organically secreted molecules to internal gene expression is referred to as 'quorum sensing' [35]. In general this refers to a switch-like change in behaviour, in response to local bacterial density. Quorum-sensing controlled responses relevant to biofilm formation include secretion of EPS [36] and modified motility [37].

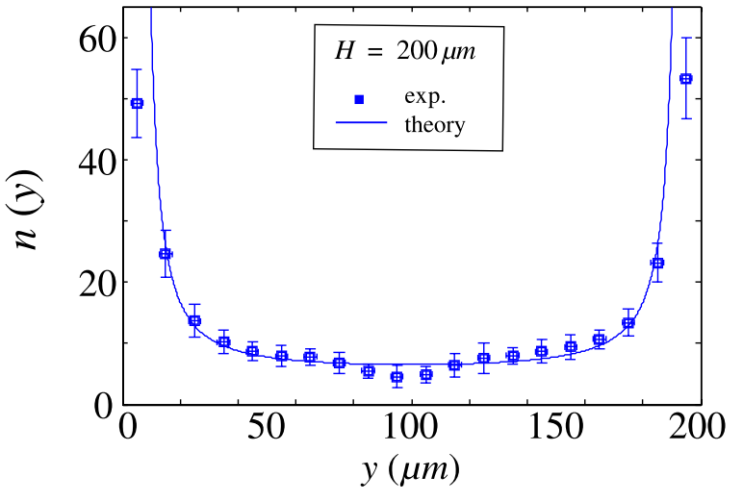


Figure 2.3 *The number density of E. coli bacteria as a function of the distance from one parallel glass plate to another, in the experiment of Berke et al., [39]. The second plate is* 200 µm *from the first. The density close the plates is at least 8 times greater than far from them, despite the lack of chemical or temperature gradients in the system.*

2.3 Micro-swimmers near surfaces

As both biofilms (and therefore surfaces) and swimming are important aspects of bacterial behaviour, their interaction is also worthy of study. It was shown by Rothschild in 1963 that bull spermatozoa, when placed between glass plates, accumulate near the surfaces [38]. A uniform density was found away from the plates, with a sharp peak at the walls. It was hypothesised that this was due to hydrodynamic effects arising from the sperm's swimming motion. More recently, it has been shown by Berke, Turner, Berg and Lauga that qualitatively similar distributions are produced by *E. coli* in a similar experiment with confinement between parallel plates [39] (Fig. 2.3). This was again explained by hydrodynamic interactions between the bacteria and the surface. The same effect was found for a strain of *Caulobacter crescentus*, a monotrichous bacterium, which only swims in the forward direction (meaning that the 'reversal' step in its 'run-reverse' mechanism does not occur) [40].

In contrast to a hydrodynamic interpretation, these results have recently been shown by Li, Tang *et al.* to be consistent with a more kinematically motivated explanation, free from explicit hydrodynamics [41]. The true picture, they argue, is that the collision of a swimmer with a surface causes it to align parallel with the surface due to steric and electrostatic effects. The swimmer then continues on

this parallel path until rotation caused by thermal noise in the fluid re-orients it such that it leaves the surface. The balance of flux towards the surface due to collisional alignment, and away from it due to rotational noise, reproduces the experimental results for spermatozoa, *E. coli* and *C. crescentus*.

Leaving aside the cause of the continued parallel swimming of these swimmers close to surfaces, the reason for their *initial* re-orientation rests on the relative strength of long-range hydrodynamic forces, to those arising from short-range, steric effects. Experiments measuring the flow of fluid near *E. coli* when it is near to surfaces suggest that the impact of nearby surfaces is small compared with thermal fluctuations, suggesting that they can initially be ignored [42]. This supports the kinematic model of Li and Tang. It should be noted that these experiments do not discount the contribution of hydrodynamics once a parallel swimming pattern have been achieved, allowing for the possibility that the true picture is somewhere between a purely hydrodynamic or kinematic explanation.

More recent work has complicated this picture further in the case of run-and-tumble bacteria such as *E. coli*, by suggesting that the bacterial swimming pattern can itself by affected by the presence of nearby boundaries. Molaei, Stocker *et al.* showed that the rate of tumbling (the amount of time spent in the 'tumbling' step outlined in Section 2.1) is decreased within 20 µm of a surface, which would provide an additional cause of such swimmers remaining near to surfaces [43]. Their suggestion for the cause of this is that the surface reduces the strength of hydrodynamic forces responsible for the flagellar unbundling that leads to the tumbling state.

From a physical background, it might be expected that such hydrodynamic interactions are largely irrelevant for the distribution of particles, as in most physical contexts, thermodynamics implies that such detailed interactions wash out at long times, leading to a Boltzmann distribution of particles. The reason such interactions *are* in fact important is that bacterial dynamics are non-time-reversible. It has been shown that a requirement for a non-Boltzmann distribution at steady-state is to have both spatial symmetry breaking (which in the parallel-plate experiment is provided by the walls) and temporal symmetry breaking (which for the micro-swimmers we have been considering is provided by their aligning interaction with surfaces) [44, 45]. The reason the aligning interaction is not time-reversible is illustrated in Fig. 2.4.

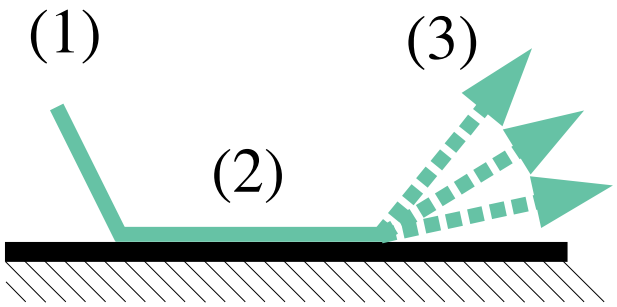
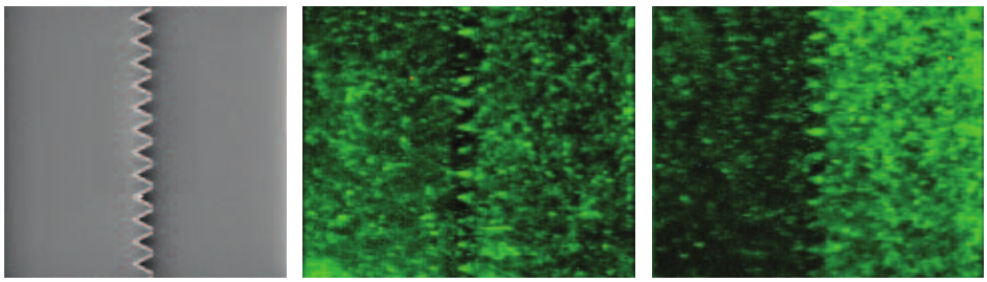


Figure 2.4 *A typical trajectory of a swimmer that aligns parallel to a surface. The incoming change in trajectory, (1) → (2), is, to a first approximation, deterministic, governed by hydrodynamic and/or steric interactions. The swimmer then leaves the surface at some later time due to stochastic rotation, (3) (this may be due to internally generated causes, for example a tumbling event, or thermal fluctuations). If we reverse the steps, the transition probability from (2) → (1) is less than unity (as would be expected for the deterministic forward process), requiring rotational noise to cause rotation to a particular orientation at a particular time. In this sense, alignment is not time-reversible.*

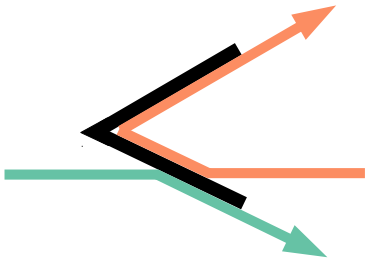
2.4 Influence of environmental geometry on bacterial distributions

In addition to the density variations induced by a planar surface, more complex distributions of surface-aligning bacteria can be achieved simply by control of the environment. A striking example of this has been shown by Galajda *et al.*, who used microfabricated 'funnels' to concentrate a population into one side of an otherwise entirely uniform chamber [46], as shown in Fig. 2.5a. The alignment effect means that an initially left-moving bacterium encountering a funnel is guided back the way it came, while an initially right-moving bacterium maintains its direction (Fig. 2.5b). This led to a density increase of three times the initial density after around an hour.

A second consequence of time-irreversible motion is the possibility of circulating fluxes in steady-state [47]. This has been exploited by Di Leonardo, Angelani *et al.* [20] (using *E. coli*) and Sokolov *et al.* [48] (using *B. subtilis*) to produce 'bacterial motors', asymmetric gears which, when immersed in a suspension of bacteria, rotate at a constant rate so long as the bacteria remain motile, driven by bacterial collisions (Fig. 2.6). This system acts to convert chemical energy in the



(a)



(b)

Figure 2.5 *(a) Creation of a non-uniform bacterial distribution through environmental topology, adapted from [46]. Left: Scanning electron micrograph of an array of funnels before loading with bacteria. Centre and right: Epifluorescence images of E. coli immediately after loading and after 80 minutes, respectively. At the later time, the density in the right side of the chamber is around 3 times the initial density. (b) A schematic illustrating the cause of the unequal distribution at steady state.*

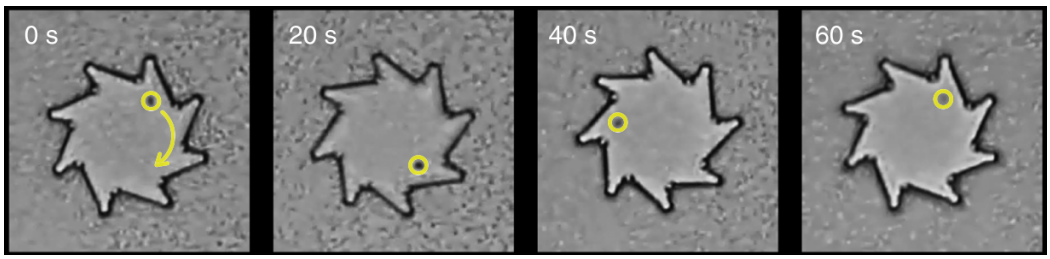


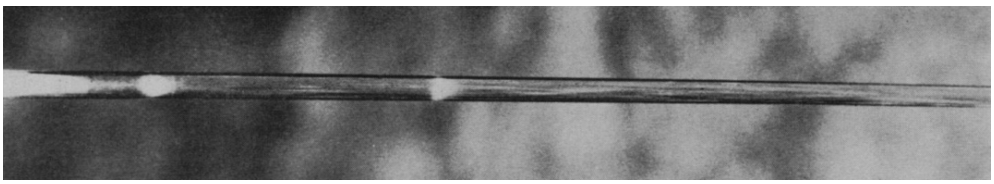
Figure 2.6 *Snapshots of a* 50 µm*-diameter gear immersed in a suspension of E. coli, from [20]. Bacterial alignment with the gear's surfaces on collision with them results in a transfer of momentum, causing a net clockwise rotation of the gear at around 1 revolution per minute.*

form of food, to mechanical work, at small scales. Angelani and Di Leonardo also developed a numerical model describing this behaviour, and used it to demonstrate an extension of the concept to translational rather than rotational motion, using what they refer to as bacterial 'shuttles' [49] (these are essentially moveable versions of the funnels used in the work by Galajda *et al.*).

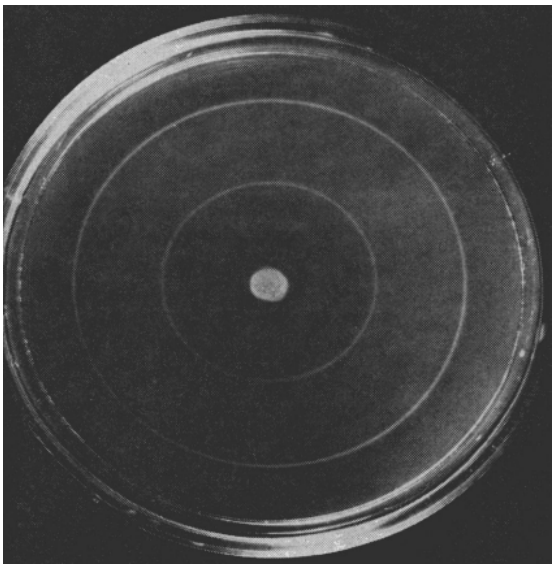
2.5 Bacterial chemotaxis

We have so far discussed the behaviour of smaller swimmers in chemically uniform environments. In the following sections we outline the behaviour of a subset of bacteria, in the situation where chemicals about which they care (such as food or poisons) are more likely to be encountered in some places than others. The ability to act to increase the probability of being in chemically favourable locations is referred to as 'chemotaxis'. Chemotaxis in bacteria was first observed in the late 19^{th} century. In 1881, Engelmann noted that a species of bacterium exhibited a repulsive response to high concentrations of carbon dioxide, performing a 'backing-up' motion which led to their expulsion from the regions of high concentrations [50]. In 1884, Pfeffer carried out further investigations by placing a capillary tube filled with chemical solutions into dilute bacterial suspensions (various species were tested), and noted accumulation at the mouth of the tube [51]. He coined the term 'chemotaxis' to describe this behaviour [52]. The suffix 'taxis' implies an active steering mechanism that is not in fact present in bacteria (explained in Section 2.5.2 below), however the term has remained in use.

In 1966 Adler demonstrated this effect using a more precise protocol, by placing a dense suspension of *E. coli* at one end of a closed capillary tube filled with uniform concentrations of a nutrient and oxygen [53, 54]. He observed multiple



(a)



(b)

Figure 2.7 *Waves in the density of E. coli arising from chemotaxis towards gradients of nutrients and oxygen created through consumption by the bacteria, from [53]. (a) Bands emerging from an inoculum at the left end of a capillary tube (b) Two rings emerging from a central inoculum, due to the consumption of two nutrient sources: glucose (consumed fastest, causing the outer ring) and galactose.*

distinct bands of cells moving to the other end of the capillary tube (Fig. 2.7a), and predicted that this was a response to gradients in the concentrations of these substances produced by their consumption. He observed similar behaviour on agar plates, where sharply defined rings of bacteria spread out from a central inoculum, in a 2D analogue of the capillary tube experiment (Fig. 2.7b).

In an attempt to explain Adler's travelling band observations, Keller and Segel produced a model of chemotaxis of the form,

$$\frac{\partial c(\mathbf{r},t)}{\partial t} = \mathrm{D}_c \nabla^2 c(\mathbf{r},t) - \delta(c)\rho(\mathbf{r},t)\,, \tag{2.1}$$

$$\frac{\partial \rho(\mathbf{r},t)}{\partial t} = \nabla \cdot \big(\mathrm{D}_\rho(c)\nabla\rho(\mathbf{r},t) - \mu(c)\rho(\mathbf{r},t)\nabla c(\mathbf{r},t)\big) \tag{2.2}$$

.

where $c(\mathbf{r}, t)$ is the concentration of chemoattractant at position $\mathbf{r}$ and time t, D_c is its diffusion coefficient, $\delta(c)$ is its rate of consumption by the bacteria, $\rho(\mathbf{r}, t)$ is the density of bacteria, $D_\rho(c)$ is the bacterial diffusion coefficient and $\mu(c)$ is the magnitude of chemotactic drift [55]. The first term on the right hand side of Eq. (2.1) represents diffusion of the chemoattractant; the second its consumption by bacteria. Eq. (2.2) is the drift-diffusion equation with diffusion coefficient D_ρ and drift coefficient μ [56]. These equations are often referred to as the 'classical' model of chemotaxis [57], and have been much studied and adapted to reproduce diverse experimental results [58, 59].

2.5.1 Secreted chemoattractants

The majority of work on chemotaxis has focused on its use to navigate in a largely static environment with stable, nutrient-rich regions [60, 61]. Budrene and Berg demonstrated that an array of complex patterns may be formed when the chemoattractant is secreted by the bacteria themselves, rather than introduced externally [62]. The experimental set-up consisted of a central inoculum of *E. coli* in a Petri dish containing a minimal growth medium, and a variety of chemicals which participate in the citric acid cycle, including succinate. This gave rise to a variety of complex patterns as shown in Fig. 2.8, which breaks the rotational symmetry of the environment, suggestive of a non-linear feedback mechanism. The cause of this was hypothesised to be metabolism of the citric acid intermediates into a chemoattractant (later identified as aspartate [63]) by the bacteria themselves, leading to a positive feedback loop where increased bacterial density at a point leads to chemical gradients, which attracts more bacteria to the point, and so on. Similar results were found for *S. Typhimirium*, separately by Woodward *et al.* [64] and Blat and Eisenbach [65].

Park, Austin *et al.* suggested that the patterns generated by Budrene and Berg were not in fact primarily due to this feedback mechanism; rather, they were the result of gradients in the chemoattractant precursor succinate [66]. This is plausible as succinate acts as a chemoattractant itself, albeit one much weaker than aspartate [67]. However Park *et al.* conducted experiments which also appear to demonstrate patterns due to self-generated chemical gradients, so it is likely the basic phenomenon is nevertheless present in some form.

These experiments involved placing uniform suspensions of *E. coli* in microfabricated, quasi-two-dimensional environments with complex geometries, specifically:

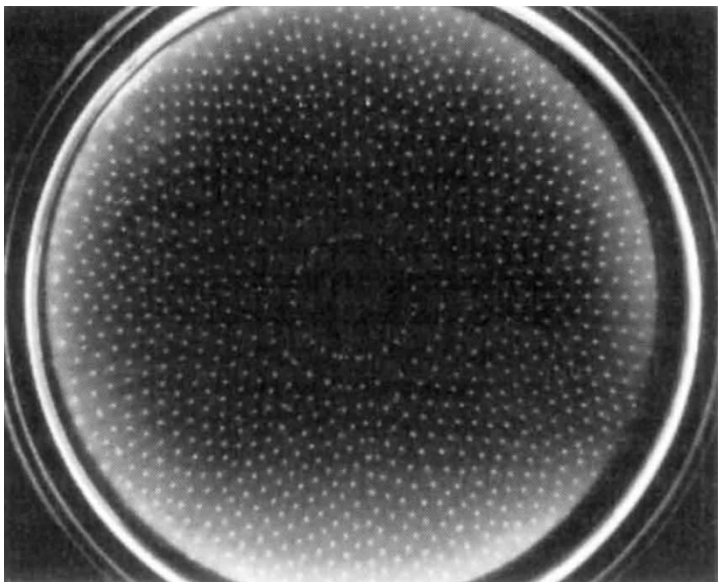
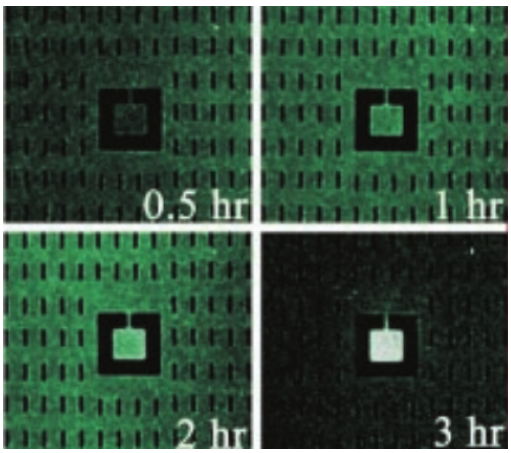


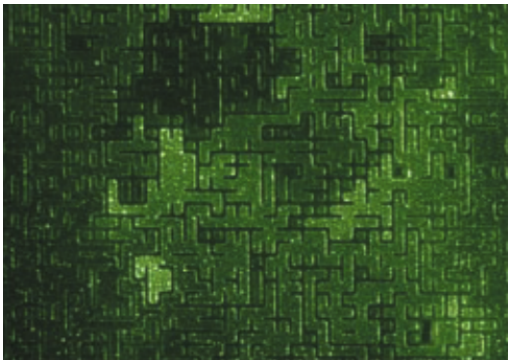
Figure 2.8 *Clustering patterns generated by E. coli from [62]. A central inoculum of cells are grown on a Petri dish, with a substrate of succinate, towards which the bacteria have a weak chemotactic response. Succinate is involved in a metabolic process resulting in the secretion of aspartate, a strong chemoattractant.*

a rectangular trap with narrow openings to allow bacteria to enter; an array of many such traps; and a randomly generated maze. The result after several hours in all cases was a non-uniform distribution of bacteria (Fig. 2.9). Replacing the bacteria with motile, non-chemotactic strains destroyed this effect (the fact that the bacteria were motile rules out the possibility that the effect was solely due to the surface interactions described in Section 2.4). The authors suggest that this clustering is evidence of bacteria using chemotaxis not only to generate dense, coordinated patterns, but also of using this to search their environment actively for places that allow for the persistence of these clusters.

This non-linear, positive-feedback chemotaxis is referred to as 'autochemotaxis' [69, 70]. It may be compared with the quorum-sensing behaviour described in Section 2.2. Quorum sensing uses chemical signalling to lower the threshold density required for biofilms to form. It is plausible that an evolutionary advantage of autochemotaxis is that it allows for the active generation of sufficient densities for biofilms to form, by attracting bacteria to one another.



(a)



(b)

Figure 2.9 *Epifluorescence images of E. coli labelled with green fluorescent protein, from [66, 68]. (a) Accumulation in a* $250\,\mu\text{m} \times 250\,\mu\text{m}$ *trap from an initially uniform distribution. At 3 hours the trap density is more than seven times than that outside of the trap. (b) Accumulation in dead ends of a random maze after 3 hours from an initial uniform distribution (not shown).*

2.5.2 Mechanism of chemotaxis

The method by which peritrichous bacteria such as *Salmonella* or *E. coli* achieve chemotaxis was established by Berg in the 1970s [71]. By tracking the paths of *E. coli* over many seconds, statistics were obtained on the distribution of 'run' and 'tumble' states as a function of the swimming direction. It was found that the average length of a tumble was constant at around 0.1 s, and in free space a run lasted around 1 s. In gradients of the amino acids serine and asparate, runs directed towards increasing concentrations lasted longer, while runs directed towards decreasing concentrations were approximately unchanged. It was therefore concluded that *E. coli* achieves a net drift towards favourable chemicals by suppressing the probability of transitioning to a tumbling state, when travelling in a favourable direction. This mechanism was later confirmed for other bacteria such as *Salmonella* and *B. subtilis* [72], although it is not universal among chemotactic bacteria: for example, *Sinorhizobium meliloti* changes its swimming speed in response to oxygen gradients [73].

2.6 *E. coli*

We conclude by re-iterating the key observations that we have described above for the sake of clarity, with emphasis on those relevant to *E. coli*, the organism we will focus on throughout the work presented here. *E. coli* is a rod-shaped bacterium, with a body roughly 2.5 µm long and 0.8 µm wide [52]. It is peritrichously flagellated, with typically six flagella protruding from its body at random points, which are 10 µm to 20 µm long [74].

E. coli moves by alternating between two states, firstly a 'run': all flagella rotate counter-clockwise, form a coherent bundle, and propel the body in a straight path; secondly a 'tumble': at least some of the flagella rotate clockwise, disrupting the bundle, and causing them to rotate independently, causing random rotation. This is called 'run-and-tumble' motion.

When the bacterium swims towards a surface, it aligns parallel with it, swims along for some time, before leaving it. This leads to accumulation near surfaces.

In the presence of gradients of chemicals for which the bacterium has suitable receptors on its surface, its distribution of states is varied so that the 'run' state is more likely when the bacterium is moving in a positive direction (towards attractants or away from repellents). The expected corresponding response to negative directions (reducing the time spent in the 'run' state) is not observed to a significant extent.

Chapter 3

Methods

3.1 Noise at small scales

When an object is placed in a fluid (a liquid or a gas) composed of much smaller particles, those constituent particles of the fluid, which are in constant motion, hit the object and induce forces on it [75]. From a macroscopic perspective, this appears as the object moving about randomly. This phenomenon is referred to as 'Brownian motion' [76]. Over a span of time much longer than those of the individual collisions, these forces can be approximated as a single force acting in a random direction. It is expected that, since the correlation times of the fluid's particles are very short, this force should be entirely uncorrelated between distinct spans of time. A second force that is relevant for small objects in a fluid is that of viscous drag, which at the small scales with which we are concerned has the form $\mathbf{F} \propto -\mathbf{v}$ (Stokes' law) [77]. Together, these give the dynamics of such an object [78],

$$m\frac{\partial \mathbf{v}(t)}{\partial t} = -\gamma \mathbf{v}(t) + \sqrt{2q}\boldsymbol{\zeta}(t)\,, \tag{3.1}$$

where γ represents the strength of viscous drag, and q the strength of thermal fluctuations. $\boldsymbol{\zeta}$ is a vector of random numbers picked from a normal distribution of zero mean and unit variance. A normal distribution is appropriate given that this is the resultant distribution of many independent random variables [79], such as kicks from many small particles. This functional form is also uncorrelated in

time, as required,

$$\begin{aligned} \langle \eta_i(t) \rangle &= 0 \\ \langle \eta_i(t) \eta_j(t') \rangle &= \delta_{ij} \delta(t - t') \,. \end{aligned} \tag{3.2}$$

As expressed here, the drag coefficient is isotropic in space, but this need not be the case in general, and its value depends on the detailed geometry of the particle.

The quantities γ and q can be related by considering the time auto-correlation of an object's velocity. Integrating Eq. (3.1),

$$\mathbf{v}(t) = \mathrm{e}^{-t/(m/\gamma)} \left(\mathbf{v}(0) + \frac{\sqrt{2q}}{m} \int_0^t \boldsymbol{\zeta}(t') \mathrm{e}^{t'/(m/\gamma)} \, \mathrm{d}t' \right) . \tag{3.3}$$

It can be seen from Eq. (3.3) that an object's velocity decays on a timescale $\tau_v = m/\gamma$. The velocity autocorrelation function can be found from Eq. (3.3), using Eq. (3.2),

$$\langle \mathbf{v}(t) \cdot \mathbf{v}(t') \rangle = \mathbf{v}(0)^2 \mathrm{e}^{-(t+t')/\tau_v} + \frac{dq}{m\gamma} \left(\mathrm{e}^{-|t-t'|/\tau_v} - \mathrm{e}^{-(t+t')/\tau_v} \right) , \tag{3.4}$$

where d is the number of spatial dimensions in the system. As the slowest decaying term, the second term dominates at long times, such that $\langle \mathbf{v}^2 \rangle = dq/(m\gamma)$. This can be compared with the average translational energy, known from the equipartition theorem [80],

$$\frac{m \langle \mathbf{v}^2 \rangle}{2} = d \frac{k_B T}{2} \,, \tag{3.5}$$

where k_B is Boltzmann's constant, to show that $q = \gamma k_B T$. It is unsurprising that the drag and noise coefficients are related, as they are both the result of random knocks from the same particles. This is an example of the fluctuation-dissipation theorem, which states that the form of fluctuations in a physical quantity *from* a source of random noise, is closely related to the dissipation of that quantity *by* that source of random noise. In this case, the physical quantity is the object's velocity.

The diffusion coefficient of such an object may be derived by considering its

free-space displacement over time, given by,

$$\mathbf{r}(t) = \mathbf{r}(0) + \int_0^t \mathbf{v}(t')\,\mathrm{d}t' \,. \tag{3.6}$$

Its mean displacement can be computed using Eq. (3.3),

$$\langle \mathbf{r}(t) - \mathbf{r}(0) \rangle = \mathbf{v}(0)\tau_v(1 - \mathrm{e}^{-t/\tau_v}) \,, \tag{3.7}$$

which becomes constant for long times. The variance of the displacement can be similarly computed,

$$\begin{aligned}
&\langle \big(\mathbf{r}(t) - \mathbf{r}(0)\big)^2 \rangle \\
=&\langle \int_0^t \mathbf{v}(t')\,\mathrm{d}t' \int_0^t \mathbf{v}(t'')\,\mathrm{d}t'' \rangle \\
=&\int_0^t \mathrm{d}t' \int_0^t \mathrm{d}t'' \langle \mathbf{v}(t') \cdot \mathbf{v}(t'') \rangle \\
=&\frac{2dk_BT\tau_v}{\gamma}\left(\frac{t}{\tau_v} - 1 + \mathrm{e}^{-t/\tau_v}\right) + \left(\mathbf{v}(0)^2 - \frac{dk_BT}{m}\right)\left(\frac{1 - \mathrm{e}^{-t/\tau_v}}{\tau_v^{-1}}\right)^2
\end{aligned} \,. \tag{3.8}$$

For $t \gg \tau_v$,

$$\langle \big(\mathbf{r}(t) - \mathbf{r}(0)\big)^2 \rangle \simeq \frac{2dk_BT}{\gamma} t \,, \tag{3.9}$$

using which the diffusion coefficient can be found from its definition,

$$\begin{aligned}
\mathrm{D} &= \lim_{t\to\infty} \frac{\langle \big(\mathbf{r}(t) - \mathbf{r}(0)\big)^2 \rangle}{2dt} \\
&= \frac{k_BT}{\gamma}
\end{aligned} \,. \tag{3.10}$$

This allows the diffusion coefficient of an object to be inferred solely from knowledge of its shape (and the temperature of the fluid).

3.2 Motion at low speeds

To find τ_v for a typical bacterium moving through a fluid, we consider it to be approximated by a sphere of an appropriate volume. For spheres, the drag coefficient is given by $\gamma = 6\pi\eta a$ [81], where a is the radius and η is the fluid's dynamic viscosity (an entirely different quantity to the thermal noise term in the previous section, despite their clashing symbols) [82]. Using $\tau_v = m/\gamma$, this means that,

$$\tau_v = \frac{m}{\gamma} = \frac{2a^2\rho_w}{9\eta}\,. \tag{3.11}$$

Bacterial densities are approximately that of water, $\rho \simeq 1 \times 10^3\,\mathrm{kg/m^3}$ [83]. We can assume the bacterium is in an aqueous solution, $\eta \simeq 1 \times 10^{-3}\,\mathrm{Pa\,s}$, and a radius of $a = 1\,\mu\mathrm{m}$ is sensible, meaning $\tau_v \simeq 2 \times 10^{-7}\,\mathrm{s}$. This implies that, for a typical bacterial speed of $v \simeq 10\,\mu\mathrm{m\,s^{-1}}$, a bacterium that stopped swimming would coast for $2 \times 10^{-3}\,\mathrm{nm}$ — around a millionth of its length — before stopping.

As an aside for the purposes of later comparison, Eq. (3.10) allows us to calculate the rate of diffusion of such a swimmer at 300 K: $\mathrm{D} \simeq 0.2\,\mu\mathrm{m^2/s}$.

As verification that inertia is unimportant, a dimensionless number which explicitly characterises the ratio of inertial forces (those which accelerate masses) to viscous forces (those opposing shear flows) is the Reynolds number, defined as,

$$\mathrm{Re} = \frac{vL\rho}{\eta}\,, \tag{3.12}$$

where v is the velocity of the fluid near the moving object, L is the scale of the motion (the size of the object) and ρ is the fluid's density [84]. For the sphere considered above, $\mathrm{Re} \simeq 1 \times 10^{-5}$. Such a small number confirms that inertia can safely be neglected for such objects [3].

The impact of this conclusion is that the acceleration term in Eq. (3.1) can be set to zero, meaning,

$$\mathbf{v}(t) = \frac{\sqrt{2q}}{\gamma}\boldsymbol{\zeta}(t) = \sqrt{2\mathrm{D}}\boldsymbol{\zeta}(t)\,. \tag{3.13}$$

Integrating this to find the position of an object after a time t has elapsed,

$$\begin{aligned}\mathbf{r}(t) &= \mathbf{r}(0) + \int_0^t \mathbf{v}(t')\,\mathrm{d}t' \\ &= \mathbf{r}(0) + \sqrt{2\mathrm{D}}\int_0^t \boldsymbol{\zeta}(t')\,\mathrm{d}t'\,, \\ &= \mathbf{r}(0) + \sqrt{2\mathrm{D}t}\boldsymbol{\zeta}(t)\,,\end{aligned} \tag{3.14}$$

There is in fact some complication to the last step of integrating the noise terms, dependent on how the term is interpreted. If the noise term $\boldsymbol{\zeta}(t)$ in the integral is considered to represent the noise as evaluated at time t (at the 'early end' of the time-step), different conclusions are arrived at than if it is considered to represent the noise half-way through the time-step, at $t + \mathrm{d}t/2$.

3.3 Swimming at low Reynolds number

One notable feature of the above analysis is the lack of a source term for the object's velocity. For passive examples such as colloids, this is the relevant case, however we are concerned with objects of a similar size, but which generate additional forces through internal means. An 'active' object commonly describes something that has an internal capacity for energy generation, storage and transformation, that is treated as a 'black box' for the purposes of analysis and simulation [85–87]. The relevant outcome of these internal dynamics is that such an object may achieve a non-zero average velocity over long times, without the presence of external fields. In the case of bacteria, this internal capacity is that of the intake of chemical energy (food), its storage, metabolism and, ultimately, conversion into mechanical work, to propel it through its environment.

The active particles we consider here are bacteria that have a switch-like swimming behaviour, which does not depend on smooth modulation of their speed. As such, we consider particles that, at a given moment, move either at a constant speed, or are entirely stopped. For the reasons outlined in Section 3.2, acceleration between these two states is not important. The remaining important variables are the distribution of the two states (moving, and stationary), and the other component of velocity: their orientation.

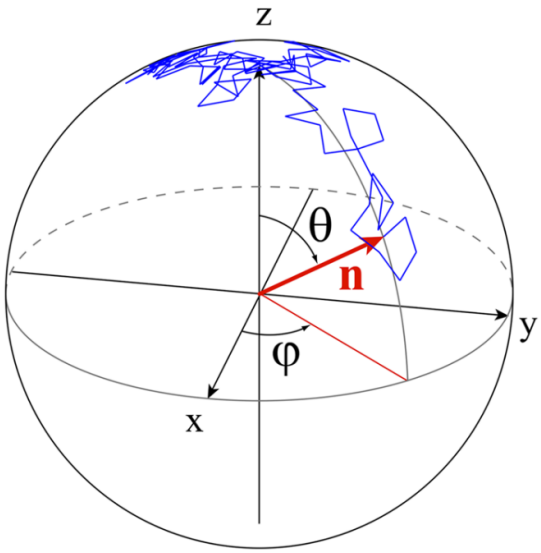


Figure 3.1 *An illustration, adapted from [88], of the diffusion of a particle's direction, initially shown by the vector* **n** *(red), in three dimensions. Its subsequent evolution is shown in blue. The analogy with a random walk on the surface of a sphere can be seen. Strictly, the periodicity of the spherical coordinates is important, however, for small rates of rotational diffusion, or early times, the assumption that the sphere is locally planar is appropriate.*

3.3.1 Brownian rotation

In addition to displacing the position of the particle in a fluid, random collisions with the fluid's constituent particles will cause random torques to be felt. For a spherically symmetric particle this makes no observable difference, however if the particle has an anisotropic shape (such as a rod-like bacterium), or a behaviour which is derived from some polarity on an otherwise symmetric body (such as an organism with a swimming direction), this rotation is an important dynamical process.

The number of translational degrees of freedom of a rigid body in d dimensions is simply the number of independent spatial directions: d. The number of rotational degrees of freedom is instead given by $d(d-1)/2$ [89]. This means that in two dimensions, there is one rotational angle; in three dimensions, there are three. So if we pick any point on the surface of a rigid body, we should expect the total angular variance to grow with time as $\langle \sigma^2 \rangle = d(d-1)\mathrm{D}_r t$, where σ is the great-circle angular distance of the point and its original position; that is, the central angle of the great circle that passes through both points. D_r is the 'rotational diffusion constant', with units of $\mathrm{rad}^2/\mathrm{s}$ since, unlike in the translational case, we

are considering displacements of angle, rather than length.

Some of these degrees of freedom represent rotations of the body about its orientational axis (for example, in three dimensions, the 'roll' component of the description of an aircraft's orientation). A simplification can be made when the diffusing object has rotational symmetry about this orientational axis. It can be seen that the arrow in Fig. 3.1 could be rotated about its own axis with no observable change (this is reflected in the fact that its location on the sphere can be specified by only two variables, here θ and φ, rather than three). For such objects, their orientation is entirely specified by a point on a $(d-1)$-sphere (a circle in 2-dimensional space, a sphere in 3). We can then study the evolution of this point as it is perturbed by collisions, by viewing it as translational diffusion on this $(d-1)$-dimensional surface (Fig. 3.1). Hence, for such objects, the number of rotational degrees of freedom is $d-1$.

Modelling this evolution through a differential equation derived by similar arguments to those in Section 3.2, we consider the evolution of a vector $\hat{\mathbf{u}}$ indicating a point on the particle's surface. Its probability density with respect to the inclination angle θ, ρ_θ, evolves as,

$$\frac{\partial \rho_\theta}{\partial t} = \mathrm{D}_r \frac{\partial^2 \rho_\theta}{\partial \theta^2} \,. \tag{3.15}$$

The same equation applies for the azimuth angle φ

By analogy with translation diffusion, to model its rotational equivalent it is sufficient to rotate $\hat{\mathbf{u}}$ by an angle picked from a normal distribution with variance $2\mathrm{D}_r\Delta t$, where Δt is the finite time-step, around a number of independent axes equal to the number of degrees of freedom, $d-1$. In two dimensions this involves a single operation,

$$\hat{\mathbf{u}}(t+\Delta t) = \mathbf{R}(\sqrt{2\mathrm{D}_r\Delta t}\eta)\hat{\mathbf{u}}(t)\,, \tag{3.16}$$

$$\mathbf{R}(\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{3.17}$$

where η again is a sample drawn from a normal distribution of unit variance and zero mean, and $\mathbf{R}(\theta)$ is a two-dimensional matrix which acts to rotate a vector by

an angle θ.

In three dimensions the same approach is used, but instead two rotation matrices are applied to rotate the vector about independent axes. This approach to modelling rotational diffusion is not computationally optimal due to its many calls to trigonometric functions, however it generates the correct dynamics [90], and is sufficient for our purposes. We verified our implementation of rotational diffusion by verifying that in free space the angular variance grows as $(d-1)\mathrm{D}_r t$ in both two and three dimensions.

Assuming Stokes flow, the rotational diffusion constant of an object can be found from its shape, using a relation analogous to the Stokes-Einstein relation for translational diffusion,

$$\mathrm{D}_r = k_B T/\gamma_r \,, \tag{3.18}$$

where γ_r is the rotational drag coefficient. For a sphere, $\gamma_r = 8\pi\eta a^3$ [91]. For the sphere considered above as an approximation of a bacterium, this implies $\mathrm{D}_r = 0.16\,\mathrm{rad}^2/\mathrm{s}$. Eq. (3.23) then implies that the self-propelled sphere we have been considering so far, swimming in three dimensions, randomises its direction every $\tau_r \simeq 3\,\mathrm{s}$.

3.3.2 Coarse-graining rotational diffusion into translational diffusion

Consider a particle initially moving in a direction $\hat{\mathbf{u}}$ at a speed v, undergoing rotational diffusion of its direction. Over a long enough time, its initial direction will be lost. The time travelled along one axis before its orientation will be lost is on average given by the inverse of the rotational diffusion constant, $\tau_r = 1/\mathrm{D}_r$. If we model this as a random walk in time increments of τ_r, and distance increments of the distance travelled before the particle is reoriented, $l_r = v\tau_r$, we have the relation for the probability density for the position, x, of the particle,

$$\rho(t+\tau_r, x) = \frac{1}{2}\left(\rho(t, x-l_r) + \rho(t, x+l_r)\right) \,, \tag{3.19}$$

which can be modified into a form closer to that of a differential equation,

$$\rho(t+\tau_r, x) - \rho(t, x) = \frac{1}{2}\left(\rho(t, x-l_r) + \rho(t, x+l_r) - 2\rho(t, x)\right) , \tag{3.20}$$

which can then be rescaled in time and space to infinitesimal quantities,

$$\tau_r\left(\rho(t+\mathrm{d}t, x) - \rho(t, x)\right) = \frac{l_r^2}{2}\left(\rho(t, x-\mathrm{d}x) + \rho(t, x+\mathrm{d}x) - 2\rho(t, x)\right) , \tag{3.21}$$

which is simply the diffusion equation,

$$\frac{\partial \rho}{\partial t} = \frac{l_r^2}{2\tau_r}\frac{\partial^2 \rho}{\partial x^2} , \tag{3.22}$$

with diffusion constant $\mathrm{D}_{T,r} = l_r^2/(2\tau_r) = v^2/(2\mathrm{D}_r)$ (the notation to make clear that this is a *translational* diffusion constant deriving from *rotational* diffusion). This is in contrast to the quantity D, which represents translational diffusion caused directly by Brownian forces from collisions with fluid particles.

To generalise this to d dimensions, we must consider the average time over which a particle's direction is randomised. In the preceding section we noted that the angular variance for a particle's direction grows as $\langle\sigma^2\rangle = d(d-1)\mathrm{D}_r t$, and for rotationally symmetric objects, we can ignore unimportant degrees of freedom so that $\langle\sigma^2\rangle = (d-1)\mathrm{D}_r t$. From this it can immediately be seen that for any degree of de-correlation we desire, the time for it to be reached is given by $t = \langle\sigma^2\rangle / \left((d-1)\mathrm{D}_r\right)$. From this, we know how the persistence time, which measures the time for a particle's direction to become de-correlated, depends on the dimensionality of the system:

$$\tau_r = 1/(d-1)\mathrm{D}_r , \tag{3.23}$$

which implies,

$$\mathrm{D}_{T,r} = v^2/(2(d-1)\mathrm{D}_r) . \tag{3.24}$$

Returning to our self-propelled sphere, this gives an effective translational diffusion constant, $\mathrm{D}_{T,r} \simeq 150\,\mu\mathrm{m}^2/\mathrm{s}$. By comparing this with its diffusion due to actual Brownian forces, $\mathrm{D} \simeq 0.2\,\mu\mathrm{m}^2/\mathrm{s}$, it can be seen that explicit Brownian diffusion is negligible compared with 'active' diffusion arising from self-propulsion.

3.3.3 Tumbling as rotational noise

In this section we map a simplified version of the run-and-tumble behaviour of bacteria such as *E. coli* onto rotational diffusion. The simple model we use is that a particle moves at a constant speed, v, with a constant orientation, $\hat{\mathbf{u}}$, during its 'run' phase. There is at all times some probability of a 'tumble' event occurring. This event is considered to be an instantaneous change of $\hat{\mathbf{u}}$, after which the run phase is resumed.

The tumble events are represented by a Poisson process, meaning that they are independent of each other: the number of tumbles in disjoint time intervals are uncorrelated. This leads to an exponential distribution of times between tumbles or, equivalently, duration of runs, τ_{run},

$$P(\tau_{\mathrm{run}}) = \alpha_0 \mathrm{e}^{-\alpha_0 \tau_{\mathrm{run}}}\,, \tag{3.25}$$

where α_0 is the 'tumble rate', with units of inverse time.

The tumbling event itself is assumed to entirely randomise the particle's direction: there is no correlation with the previous value of $\hat{\mathbf{u}}$. In reality, for *E. coli* there is in fact some correlation, with the difference between new and old angles on average 60°, rather than 90°, as would be expected for complete randomisation [92]. Further, there is evidence that this distribution of new angles is itself systematically varied by *E. coli* so as to improve their efficiency at doing chemotaxis [93]. For simplicity, however, we neglect both of these effects.

For the exponential distribution Eq. (3.25), $\bar{\tau}_{\mathrm{run}} = \alpha_0^{-1}$ [94]. From the definition of a tumbling event chosen, it is clear therefore that the average time for a particle's direction to be randomised is α_0^{-1}. This means that an effective rotational diffusion constant can be assigned to this value using Eq. (3.23),

$$\tilde{\mathrm{D}}_r = \frac{\alpha_0}{(d-1)}\,, \tag{3.26}$$

such that at long times, the displacement of a particle undergoing rotational diffusion at this rate, is identical to the displacement of a particle not doing rotational diffusion, but which tumbles at a rate α_0.

In reality, these two sources of rotational noise, Brownian rotation and tumbling, are present together. A derivation of the effective rotational diffusion constant when both are present was derived by Lovely and Dahlquist [95],

$$\tilde{\mathrm{D}}_{\mathrm{r,eff}} = \mathrm{D}_r + \frac{\alpha_0}{(d-1)}\,, \tag{3.27}$$

which gives rise to an effective translational diffusion constant,

$$\mathrm{D}_{T,\mathrm{r,eff}} = \frac{v^2}{2\left((d-1)\mathrm{D}_r + \alpha_0\right)}\,. \tag{3.28}$$

A typical rate of tumbling for *E. coli* is $\alpha_0 \simeq 1\,\mathrm{s}^{-1}$ [88]. This gives a total effective translational diffusion constant, due to swimming, of $\mathrm{D}_{T,\mathrm{r,eff}} \simeq 40\,\mu\mathrm{m}^2/\mathrm{s}$.

For the sake of completeness, an expression for the total effective translational diffusion constant, $\mathrm{D}_{\mathrm{eff}}$, combining translational and rotational diffusion and tumbling, can be expressed. The necessary additivity of terms is shown in [96],

$$\mathrm{D}_{\mathrm{eff}} = \mathrm{D} + \frac{v^2}{2\left((d-1)\mathrm{D}_r + \alpha_0\right)}\,. \tag{3.29}$$

The tumbling event has so far been described as a homogeneous Poisson process, meaning that α_0 is not a function of t. As described later, in fact we are modelling an *in*homogeneous Poisson process, with an implicit time dependence through the particle's chemotaxis machinery. In this case, the number of tumbling events between $t - \Delta t$ and t is given by,

$$N_t = \int_{t-\Delta t}^{t} \alpha(t')\,\mathrm{d}t'\,. \tag{3.30}$$

In our implementation, we assume that α varies sufficiently slowly over the interval

Δt that we can assume it is constant,

$$N_t = \alpha(t - \lambda\Delta t)\Delta t\,, \tag{3.31}$$

where λ is a parameter between 0 and 1 controlling at what point in time the noise term is evaluated. In our case we do not know $\alpha(t' > t - \Delta t)$, because in our simulations we model time as discrete, in multiples of Δt, so α is not evaluated for intermediate values. For this reason we are forced to choose $\lambda = 1$, however it should be noted that this approximation is not optimal, since it is biased towards the value of α at the early end of the interval. In any case, assuming Δt is sufficiently small, we can assume that $\alpha(t - \Delta t)\Delta t \ll 1$, and use Eq. (3.31) as the probability in our simulation that a particle should undergo a single tumbling event in a time-step of size Δt.

3.4 Numerical solutions of Partial Differential Equations (PDEs)

3.4.1 Introduction

In several of the situations that are modelled in the following chapters, there is a chemical whose concentration affects the dynamics of self-propelled particles. Since there are many more molecules of the chemical than particles of the bacteria, it is sensible to treat the chemical as forming a continuous concentration field.

The dynamics such a field undergoes depends on the particular context, but in all relevant cases, it can be separated into two parts: diffusion and kinetics. Diffusion involves the dispersal of the chemical by Brownian motion, and is governed by the diffusion equation,

$$\frac{\partial c(\mathbf{r}, t)}{\partial t} = \mathrm{D}_c \nabla^2 c(\mathbf{r}, t)\,, \tag{3.32}$$

where c represents the chemical concentration at a point $\mathbf{r}$ and time t, and D_c the diffusion constant of the chemical.

Chemical kinetics involves the chemical processes that lead to the breakdown,

or formation of the molecule. This could include secretion by an organism, or degradation by thermal agitation. These are local phenomena that do not depend on gradients in the chemical. So the dynamics of such a chemical field are of the form,

$$\frac{\partial c(\mathbf{r},t)}{\partial t} = \mathrm{D}_c \nabla^2 c(\mathbf{r},t) + f(\mathbf{r},t,c,\ldots)\,, \tag{3.33}$$

where $f(\mathbf{r},t,c,\ldots)$ is an arbitrary function representing the chemical kinetics. The implied additional parameters characterise the particular rates of the relevant mechanisms.

3.4.2 Operator splitting

When solving partial differential equations numerically, ideally the equation is of a form that is well-known, and efficient methods have already been developed. Usually this is not the case, however. In some of these difficult cases, the equation can be decomposed into a number of sub-equations, which are simpler and efficiently solvable — this leads to the method known as 'operator splitting' [97], which is outlined below. The method we explain is the lowest-order implementation, known as 'Lie-Trotter sequential splitting' [98]; more elaborate implementations have been developed to reduce the error introduced by the splitting, but this was not found to be necessary in our case.

If there is an initial value problem of the form,

$$\frac{\partial u}{\partial t} = F(u)\,, \tag{3.34}$$

where F is some operator, and if the operator can be decomposed as,

$$F = F_1 + F_2\,, \tag{3.35}$$

then operator splitting involves solving the set of equations,

$$\begin{array}{l} \frac{\partial u_1}{\partial t} = F_1(u) \\ \frac{\partial u}{\partial t} = F_2(u_1) \end{array}. \tag{3.36}$$

This approach has many advantages. The primary motivation is that the sub-equations may be of forms that are amenable to specialized, efficient algorithms. It also allows these sub-problems to be solved independently, meaning that the scheme used to solve one can be improved without having to consider the other components of the full PDE. Using the best available method to solve each component can improve the stability of the overall algorithm. It also becomes easier to add or remove terms in the PDE to modify the model, without needing to rewrite large portions of the implementation [99].

This approach is used here in solving Eq. (3.33). The method has been successfully applied to this particular class of PDEs in studying the contamination of groundwater [100–102]. Using their choice, the equation is decomposed into,

$$\begin{aligned} \tfrac{\partial c_1}{\partial t} &= \mathrm{D}_c \nabla^2 c(\mathbf{r}, t) \\ \tfrac{\partial c}{\partial t} &= f(\mathbf{r}, t, c_1, \ldots) \end{aligned} \,. \tag{3.37}$$

The first equation is the diffusion equation, which as an extremely common problem has well-known, efficient numerical solver methods. The second equation is purely local, with no dependence on spatial gradients, which means that efficient and stable schemes for solving Ordinary Differential Equations (ODEs) can be employed.

3.4.3 Numerically solving the diffusion equation

We consider first the problem of solving the diffusion equation in one dimension, for an arbitrary initial condition $c(x, t = 0) = c_0$,

$$\frac{\partial c}{\partial t} = \mathrm{D}\frac{\partial^2 c}{\partial x^2} \,. \tag{3.38}$$

The equation can be solved for any c_0 by the method of 'finite differencing'. This essentially involves replacing the infinitesimal quantities in a differential equation, in this case ∂x and ∂t, with small but finite quantities, Δx and Δt, and iteratively solving the resultant equation for a discrete set of values of (x, t) [103].

The time derivative can be approximated by,

$$\frac{\partial c}{\partial t} \simeq \mathrm{D}\frac{c(x,t+\Delta t)-c(x,t)}{\Delta t}\,. \tag{3.39}$$

There are a number of ways to make this approximation. This particular form is referred to as 'forward explicit' differencing. It is 'forward' because its approximation is weighted towards the value of the function at the later time, $t+\Delta t$. It is 'explicit' because the state of the system at $t+\Delta t$ depends only on its state at t; 'implicit' methods are possible, where the state at $t+\Delta t$ is found from an expression that involves the state at both t and $t+\Delta t$.

In order to calculate the spatial derivative, it is useful to inspect the Taylor expansion of a function about the two points in discretised space adjacent to (x,t).

$$\begin{aligned} c(x+\Delta x,t) &\simeq c(x,t)+\Delta x\tfrac{\partial c(x,t)}{\partial x}+\frac{\Delta x^2}{2}\tfrac{\partial^2 c(x,t)}{\partial x^2} \\ c(x-\Delta x,t) &\simeq c(x,t)-\Delta x\tfrac{\partial c(x,t)}{\partial x}+\frac{\Delta x^2}{2}\tfrac{\partial^2 c(x,t)}{\partial x^2} \end{aligned}\,. \tag{3.40}$$

These can be combined to give an expression for the second derivative,

$$\frac{\partial^2 c}{\partial x^2} \simeq \frac{c(x+\Delta x,t)+c(x+\Delta x,t)-2c(x,t)}{\Delta x^2}\,. \tag{3.41}$$

This approximation is not biased in any spatial direction, and so is a 'centred' difference rule. Combining Eqs. (3.39) and (3.41) gives,

$$c(x,t+\Delta t) \simeq c(x,t)+\frac{\mathrm{D}\Delta t}{\Delta x^2}\left(c(x+\Delta x,t)+c(x-\Delta x,t)-2c(x,t)\right)\,, \tag{3.42}$$

which can be directly solved to solve the diffusion equation numerically over a time Δt. It can be seen that the left-hand side only depends on the system's state at time t, showing that it is an explicit method.

It is clear that if Δx or Δt are too large, the solution will not correctly capture the dynamics of the system. The criterion under which the approximation can be expected to be stable (meaning that numerical errors do not grow exponentially over time) can be found using a von Neumann stability analysis.

This method essentially involves considering a small perturbation, ϵ, to an exact solution, $\tilde{c}$, and investigating how the perturbation evolves over time. We can write this as,

$$c(x,t) = \tilde{c}(x,t) + \epsilon(x,t)\,. \tag{3.43}$$

This can be substituted into Eq. (3.42) to find how ϵ is expected to evolve in time,

$$\epsilon(x,t+\Delta t) \simeq \frac{\mathrm{D}\Delta t}{\Delta x^2}\left(\epsilon(x+\Delta x,t) + \epsilon(x-\Delta x,t) - 2\epsilon(x,t)\right)\,. \tag{3.44}$$

Expanding ϵ into Fourier modes,

$$\epsilon(x,t) = \sum_j A_j(t) \exp ik_j x\,, \tag{3.45}$$

where $k_j = 2\pi j/L$, and L is the length of the system. We can inspect a single mode, and assume that the error from this mode either grows or decays with time,

$$\epsilon(x,t) = \lambda^t \exp ikx\,, \tag{3.46}$$

where λ indicates whether the error mode decays ($|\lambda| < 1$) or grows ($|\lambda| > 1$) in time. Substituting Eq. (3.46) into Eq. (3.44) gives,

$$\lambda = 1 + \frac{\mathrm{D}\Delta t}{\Delta x^2}\left(2\cos(k\Delta x) - 2\right)\,. \tag{3.47}$$

By noting that it is always true that $2\cos(k\Delta x) - 2 < 0$, the largest λ could be for any set of parameters is $\lambda = 1$. Therefore, the only way the error mode can be

unstable is when $\lambda < -1$. Hence, for numerical stability we require $\lambda > -1$,

$$1 + \frac{\mathrm{D}\Delta t}{\Delta x^2}\left(2\cos(k\Delta x) - 2\right) > -1\,, \tag{3.48}$$

which can be solved to give a stability criterion for this solution of the diffusion equation,

$$\frac{2\mathrm{D}\Delta t}{\Delta x^2} < 1\,. \tag{3.49}$$

The implication of this is that if we wish to double our spatial resolution, our time-step must be reduced by a factor of four.

3.4.4 Solving Ordinary Differential Equations (ODEs)

In solving the kinetic portion of the differential equations used here, the fact that they do not depend on spatial gradients means that methods used to solve ODEs can be used. It is possible to do this using the Euler method outlined above for the diffusive portion, however this is neither the most stable, nor most efficient method available. This is also true when solving the diffusion equation, where methods such as the implicit Crank-Nicolson method are superior [104]. However this is difficult to implement in two dimensions, and the complicated spatial boundary conditions introduced by an environment with a complicated geometry makes implementation unfeasible.

The Euler method evaluates the change in a function's value at a particular time-step, and extrapolates this forward to the next time-step. This approximation of the time-derivative is asymmetrically weighted towards the value of the function at t. The essential idea behind the Runge-Kutta family of methods is to evaluate the function slope at points between t and $t + \Delta t$, in order to reduce the error in the solution. There is a choice of how many midpoint evaluations to use, with increasing accuracy at the cost of an increased number of computations. The most commonly used compromise is the 'Fourth-order Runge-Kutta' method, which uses the following scheme [105]: for an ODE of the form,

$$\frac{\mathrm{d}u}{\mathrm{d}t} = f(t, y)\,, \tag{3.50}$$

$$u(t+\Delta t) = u(t) + \frac{\Delta t}{6}\left(k_1 + 2k_2 + 2k_3 + k_4\right) , \tag{3.51}$$

where,

$$\begin{aligned} k_1 &= f(t, y(t)) \\ k_2 &= f(t + \frac{\Delta t}{2}, y(t) + \frac{\Delta t}{2} k_1) \\ k_3 &= f(t + \frac{\Delta t}{2}, y(t) + \frac{\Delta t}{2} k_2) \\ k_4 &= f(t + \Delta t, y(t) + \Delta t k_3) \end{aligned}$$

.

3.5 Packing objects using the Metropolis-Hastings algorithm

In Chapters 4 and 5, we require high-density packings of geometrical objects in bounded spaces. The most naïve method is to repeatedly place all the objects at random uniform locations, and to accept a configuration when there are no intersections. This produces uniform distributions, and is sufficient for low packing fractions (the fraction of space occupied by the objects). For larger volume fractions (the threshold depends on the shape of the particular object) random placement becomes prohibitively slow. As a more efficient solution, we use a modified version of the Metropolis-Hastings algorithm, which we outline below.

The Metropolis-Hastings algorithm is a Monte Carlo method, originally developed in order to sample an unknown probability distribution [106]. This is useful when the parameter space is high-dimensional, where directly picking a sample subset of all states would be impractical [107]. As with other Markov Chain Monte-Carlo (MCMC) methods, the essence of the algorithm is a random walk between a set of states, representing the parameter space of the system in question, with transitions between states at each step determined probabilistically. In the case of the Metropolis-Hastings algorithm, the transition probability is dependent on the difference in value of the probability distribution, evaluated for the current and

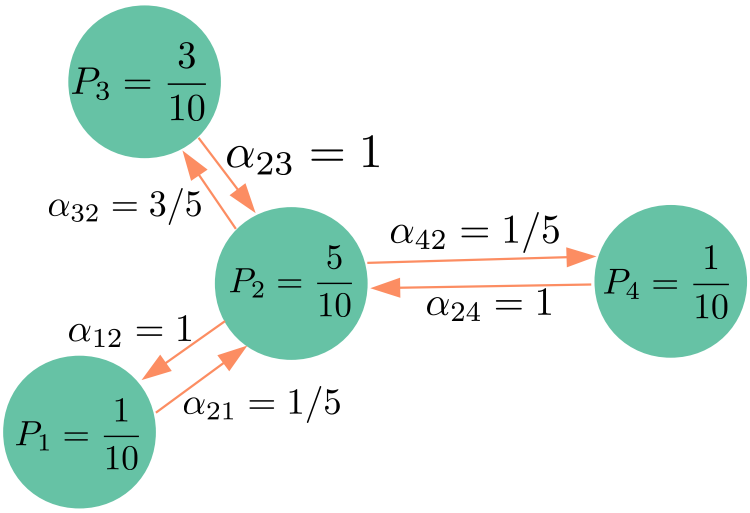


Figure 3.2 *An example of a Markov chain. In this case there are four states (green) in the chain. P_i represents the probability distribution that we wish to sample. Arrows indicate transition probabilities between these states when carrying out the Metropolis-Hastings algorithm, which are computed as follows: if the proposed next state, j has a higher value of P than the current state, i, the current state becomes that proposed state with probability 1. Otherwise, the probability of the move is proportional to the ratio of the probabilities of each state, P_j/P_i. Over long times, the fraction of time the algorithm spends in state i tends towards P_i.*

target states. By a suitable choice of this stepping probability, the fraction of time spent in each state will tend towards the distribution's probability at that point.

The essence of the method, with respect to its implementation, is summarised in the following algorithm:

current state = a random pick from **all states**
for n = 0 .. $n_{\max}$ **do**
    new state = a random pick from **adjacent states**
    α = Probability(new state) / Probability(current state)
    if $\alpha > 1$ **then**
        current state = **new state**
    else
        current state = **new state** with probability α
    end if
end for

In the original formulation, the aim was to accurately sample states in the canonical ensemble in statistical mechanics [108], also known as the *NVT*

ensemble as it describes a system with a fixed Number of particles, Volume and Temperature [109]. For a given configuration of particles, its probability of occupation is that given by statistical physics [110],

$$P(\{\mathbf{r}\}) \propto \exp(-\beta U(\{\mathbf{r}\}))\,, \tag{3.52}$$

where $\{\mathbf{r}\}$ is the set of particle coordinates, U is the energy of this new configuration (which is explicitly calculated from pairwise-potentials), and $\beta = 1/k_B T$, where k_B is Boltzmann's constant, represents the degree of thermal noise in the system. Changes between configurations are achieved by perturbing the coordinate of a particular particle in the ensemble along one axis,

$$\mathbf{r}_{i,d} \to \mathbf{r}_{i,d} + \Delta r \eta\,, \tag{3.53}$$

where i is a random, uniform choice over all particles in the system, d is a random, uniform choice over all axes in the system, η is a random number chosen from some probability distribution, and Δr is a free parameter of the algorithm, controlling the speed at which system configurations are explored. A large value causes states to be explored in a short time, but limits how well small, high-probability regions can be resolved. The opposite trade-offs arise with small values.

This approach can be easily extended to model the isothermal-isobaric, NpT ensemble, where the Volume of the system can vary but which maintains constant $p$ressure [111]. The quantity to consider in the probability calculation is now [112],

$$P(\{\mathbf{r}\}, V) \propto \exp(-\beta(U(\{\mathbf{r}\})) + pV))\,. \tag{3.54}$$

The potential states that must be explored now includes those with a range of values for V. This is accomplished within the algorithm by incorporating a probability to perturb the size of the system (rather than a particle's position),

$$L_d \to L_d + \Delta L \eta\,, \tag{3.55}$$

where L_d indicates the length of the system along one axis, and ΔL is an additional free parameter of the system playing a similar role to Δr. A common choice to decide when to perturb the system size, or when to perturb a particle coordinate,

is to perturb the size as often as a given single particle is perturbed.

With suitable modification, the NpT form of the Metropolis-Hastings algorithm can be used to quickly construct high-density packings of geometric objects inside a bounded cuboid space. In this case we do not care about sampling the 'space of possible packings' appropriately, but instead simply wish to find a packing state with the particular property we want: a given packing fraction. For geometric objects with hard surfaces, the appropriate interaction strength between two objects, i and j, is hard-core:

$$U(\mathbf{r}_i, \mathbf{r}_j) = \begin{cases} \infty & \text{object } i \text{ intersects object } j \\ 0 & \text{otherwise} \end{cases} . \tag{3.56}$$

Some sense of how the algorithm proceeds for this case is as follows: the system is initialised with a very low packing fraction, which is trivially generated by random placement of the objects in a large volume. The system's size is quickly decreased, due to the pV term in Eq. (3.54), until a decrease in the system's size would cause an object to intersect the bounding box. The objects rearrange themselves until the volume can be reduced, and this process repeats until the desired packing fraction is achieved, at which point the algorithm stops.

Despite the packing algorithm having no connection to thermodynamics, there remains the choice of β. This is now a non-physical choice that only affects the speed of convergence. In such circumstances, an additional tool may be used to improve the algorithm's efficiency, known as 'simulated annealing' [113]. The essential idea is to choose a small β (large noise) at the beginning of the algorithm, when it is important to quickly explore the state-space to find the broad region of interest. β is then smoothly increased in order to reduce the size of the region, hopefully converging on the optimal state. In our algorithm, a linear ramp of β was used, proportional to the packing fraction of the system.

3.6 Vicsek model

The Vicsek model provides a minimal description of the dynamics of interacting self-propelled particles, in such a way that 'flocking' behaviour can be observed, where there are long-range correlations in the particles' orientations. As a classic example

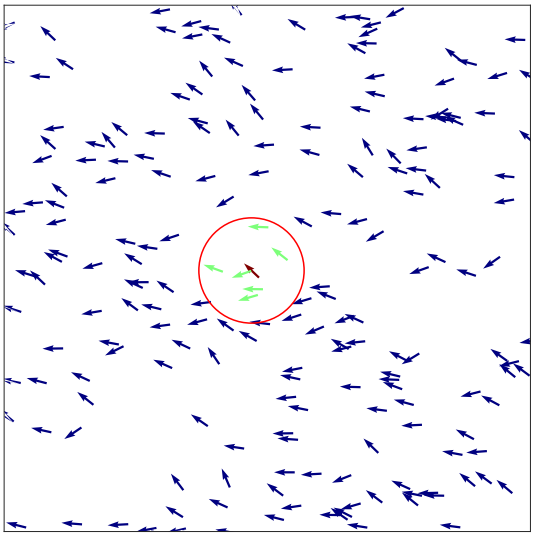


Figure 3.3 *A visual explanation of the Vicsek algorithm. For a particular particle (red arrow) at a given time-step, it considers the velocity of all its neighbours (green arrows), within a circle of radius r_v (red circle). Its velocity at the next time-step is the average of its neighbours, and itself, with some rotational noise, picked from a uniform distribution.*

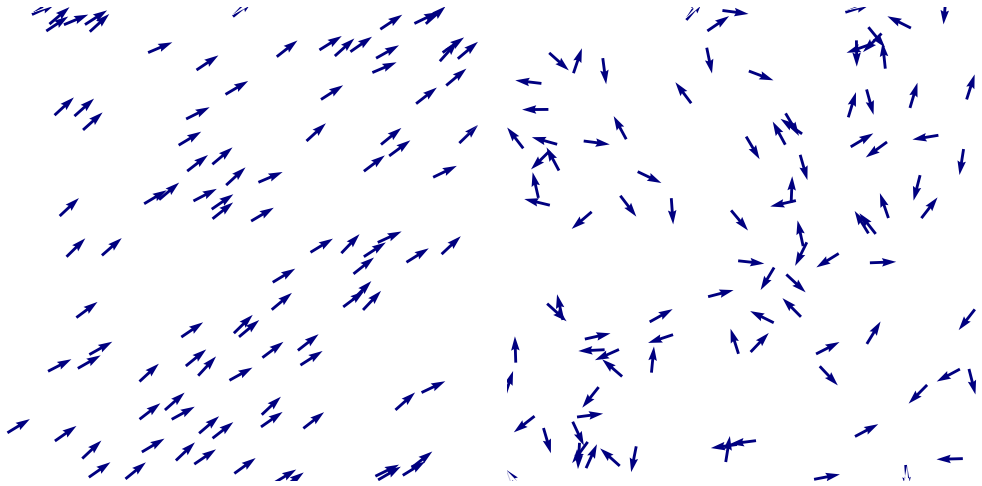
of agent-based simulations, we explore its properties below, as an introduction to the method of agent-based modelling.

In the original formulation of the model as described by Vicsek [114], there are N particles moving in a two-dimensional space of length L, with periodic boundary conditions. Each particle, i at time t has a position $\mathbf{r}_i(t)$ and a velocity $\mathbf{v}_i(t)$, whose magnitude is the same for all particles, $v_i = v$. At each timestep, each particle considers the velocities of all other particles within a distance radius r_v and aligns its own velocity according to,

$$\mathbf{v}_i(t + \Delta t) = \mathbf{R}(\xi_i(t))\frac{1}{\mathrm{N}_i}\sum_{r_{ij}<r_v}\mathbf{v}_j(t)\,, \tag{3.57}$$

where $\mathbf{R}$ is a matrix which acts to rotate a vector by an angle $\xi_i(t)$, which is chosen from a uniform distribution between $-\eta$ and η, N_i is the number of neighbours within a distance r_v, and the sum is over all particles within that radius Fig. 3.3.

The parameters of the model are therefore the density of particles in the system, $\rho = \mathrm{N}/\mathrm{L}^2$, and the amount of noise in the system, η. At low η, the system exhibits polar order, while at sufficiently high η, this order is lost (Fig. 3.4). The typical order parameter used in measuring this behaviour is the magnitude of the average



(a) **(b)**

Figure 3.4 *Snaphots of the (a) ordered and (b) disordered states of the Vicsek model.*

macroscopic velocity,

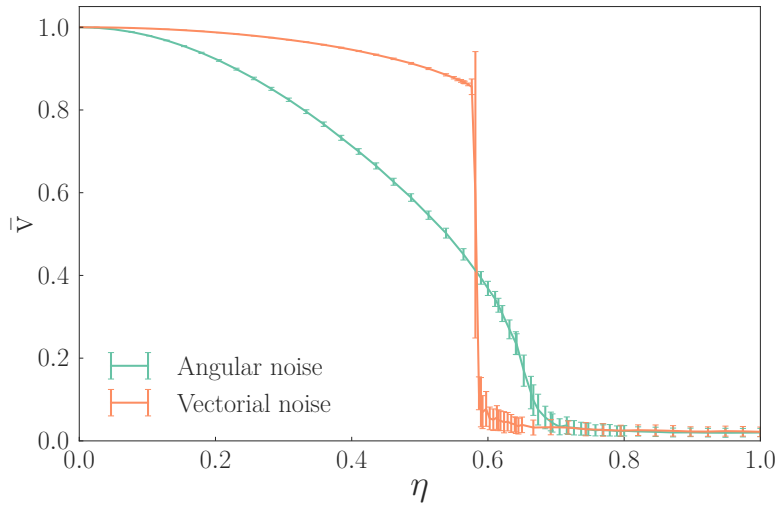
$$\bar{v}(t) = \frac{1}{\mathrm{N}v}\left|\sum_i \mathbf{v}_i(t)\right| . \tag{3.58}$$

There is a phase transition between a flocking and ordered state, as shown by the change in the order parameter, and an increase in the size of its fluctuations at the critical point (Fig. 3.5). The smooth variation of the order parameter and its variance implies a second-order phase transition, at $\eta \simeq 0.66\,\mathrm{rad}$. A variant of Eq. (3.57) has also been studied by Chaté *et al.* which implements noise differently. The noise in Eq. (3.57) may be seen as an 'alignment error', as it is applied once to each particle to the true average neighbourhood velocity. If the noise is seen as reflecting 'measurement error', then the random noise should be applied to the velocity of each neighbouring particles. This leads to the rule,

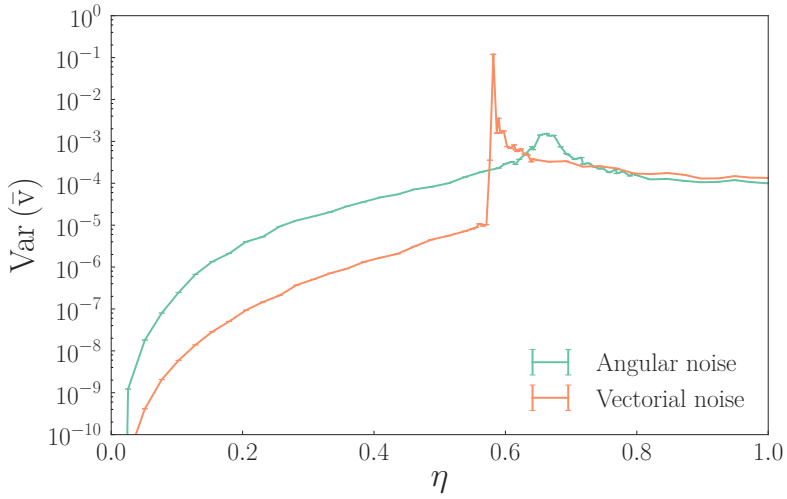
$$\mathbf{v}_i(t+\Delta t) = \frac{1}{\mathrm{N}_i}\sum_{r_{ij}<r_v} \mathbf{R}(\xi_{ij}(t))\mathbf{v}_j(t) . \tag{3.59}$$

The original Vicsek formulation (alignment error) has been referred to as 'angular noise', while the later Chaté model (measurement error) has been referred to as 'vectorial noise'. In this model, the nature of the transition changes from second- to first-order, as shown by a discontinuity in the order parameter (Fig. 3.5). There

has been some controversy about the assignation of these orders to each of these models, as the model is sensitive to seemingly minor detail such as the order in which positions and orientations are updated, and to finite-size effects. Thorough investigation by Baglietto and Albano [115] has shown that the orders stated are the correct ones, when the system's boundaries are changed dynamically during the simulation to avoid artificial symmetry breaking.



(a)



(b)

Figure 3.5 *Macroscopic metrics of a set of particles implementing some variant of the Vicsek model, as a function of its noise parameter, η. We have gathered data by directly simulating the Vicsek model for the given parameters. The parameters used were* $\mathrm{N} = 2048$, $\mathrm{L} = 32$, $v = 0.5$, *where lengths are in multiples of the alignment radius* r_v. *The angular noise model (green) interprets the noise parameter as parametrising an imperfect attempt by each particle to align with the well-measured average neighbourhood direction. The vectorial noise model (orange) interprets the noise as parametrising each particle's perfect alignment with an imperfectly measured average neighbourhood direction. (a) The magnitude of the net velocity of all particles, normalised by their speed. The angular noise variant shows a a smooth transition, and no obvious location of a transition. The vectorial noise variant shows strong evidence of a discontinuity at* $\eta \simeq 0.58$, *implying a first-order phase transition. (b) The variance in the order parameter over successive configuration snapshots after the system has reached a steady state. A transition point in the angular noise variant can be identified at* $\eta \simeq 0.66$, *implying a second-order phase transition.*

Chapter 4

Steady-state distributions of motile *E. coli* in water-in-oil emulsions

4.1 Introduction

The behaviour of self-propelled particles in confined spaces is an area of statistical physics that is commonly probed using motile bacteria. They have the necessary ingredients of active propulsion, simple geometry, and minimal behaviour (compared with more complex biological agents such as fish or wildebeest).

In experiments where swimmers accumulate at the edges of parallel glass plates, such as [38, 39, 41], there is strong confinement along the axis perpendicular to the glass slides, but none along the other axes. In these cases there is a natural direction along which accumulation of bacteria occurs — the one perpendicular to the glass slides.

More recently, confinement of *Bacillus subtilis* within squashed water droplets was demonstrated. Here, there is planar confinement along one axis as in the examples above, and also confinement with two-dimensional rotational symmetry [116] along the others (see Fig. 4.1). It was found that a layer of bacteria is formed at the droplet's curved interface, whose members orbit coherently in a single direction, while the remaining bacteria coherently rotate in the same plane, in the *opposite* direction. In this case also, there is a natural direction along which the

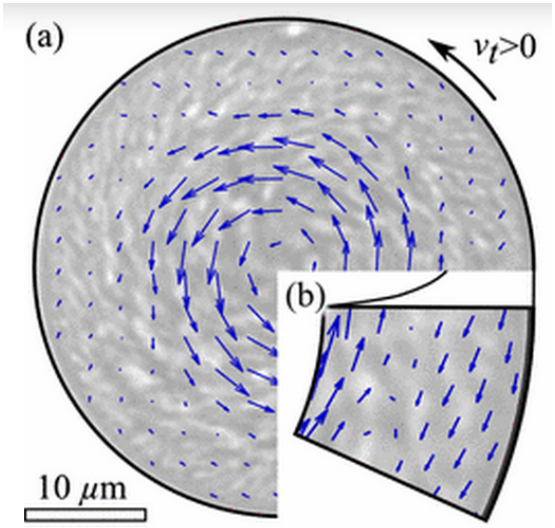


Figure 4.1 *The fluid flow field at steady-state for a cross-section through a droplet filled with a dense suspension of Bacillus subtilis, from [116]. (a) The flow over the entire cross-section. (b) A region close to the interface, showing a thin counter-rotating layer. The volume fraction occupied by the bacteria is* 40 %.

accumulation occurs — that parallel to the glass slides.

An interesting case remains, where the confining space is completely isotropic. In this case there is no particular direction in which bacteria can stably align, and complex phenomena may be expected. In practice this means that the bacteria are in a sphere, whose radius is at most on the order of the bacterium's persistence length. This arrangement is the one investigated here.

Such a situation is commonly realised in nature. *Sinorhizobium meliloti* fixes atmospheric nitrogen into ammonium, a more biologically useful form [117]. *S. melitoti* forms a symbiotic relationship living in root nodules of legumes such as Alfalfa. From water droplets or films on root hairs, the bacteria form a biofilm that causes the host to form such root nodules. Rain droplets are more generally important in the dispersal of plant pathogens [118]. Such pathogens include the motile organism *Xanthomonas campestris* [119, 120].

Emulsions with water droplets with a size less than the persistence length of motile bacteria, (µms) may be useful environments in which to host microbial cells for research purposes [121]. Confining microbes in such emulsion droplets may also be used to create a large number of independent populations on which to do measurements [122]. It has been shown that when *Bacillus subtilis* are introduced to the gastrointestinal tracts of chickens, where they act as a probiotic

by outcompeting *Campylobacter jejuni*, those with higher motility are more effective in reducing the incidence of gastroenteritis [123]. Microspheres of similar diameter to the droplets studied here are of interest in controlling the delivery of drugs and nutrients to humans [124].

4.2 Experimental setup

The simulations carried out here were done simultaneously with experimental work carried out by I. D. Vladescu, studying spherical water droplets suspended in oil. We outline the experimental arrangement below for the sake of providing context. The emulsion on which imaging was done consisted of sunflower oil, in which was dispersed a solution of *Escherichia coli*. The solution was an aqueous phosphate buffer, which is non-toxic to the bacteria but does not lead to bacterial growth [125], containing bacteria at cell densities of (1.55 to 4.65) $\times 10^9\,\mathrm{ml}^{-1}$. We refer to droplets of this aqueous solution as 'water droplets' hereafter for the sake of simplicity. The bacteria were of strain AB1157, and had the gene encoding for the signalling protein 'CheY' deleted, which causes the bacteria to no longer tumble [126]. The strain was also modified to express green fluorescent protein (GFP), which is a method commonly used to make the bacteria more easily visible so that their positions can be determined [127].

These water droplets were then imaged using high-resolution fluorescent confocal microscopy, from which the positions of the individual bacteria were recorded. Fig. 4.2 shows 2 µm-thick cross-sectional images of the equator of typical droplets from these experiments. A single such stack of images representing a droplet consisted of 5 to 60 cross-sectional images at varying depths along the imaging axis. The distribution of the bacteria's speeds was also recorded using differential dynamic microscopy (DDM) [128, 129]. This distribution was narrow enough to be assumed constant in subsequent analysis. The imaging process was carried out ten times for each droplet; that is to say, there were ten images at a particular distance along the imaging axis.

This data was collected for a range of droplet radii, R, which spanned $R = 5\,\mu\mathrm{m}$ to $40\,\mu\mathrm{m}$, with most in the range $R = 10\,\mu\mathrm{m}$ to $20\,\mu\mathrm{m}$. Data was also collected for a range of bacterial densities, ρ_0. For ease of interpretation this is typically quoted below as the fraction of the droplet's volume occupied by the bacteria, ϕ. The two are simply related by $\phi = \rho_0 V_b$, where V_b is the volume occupied by a single

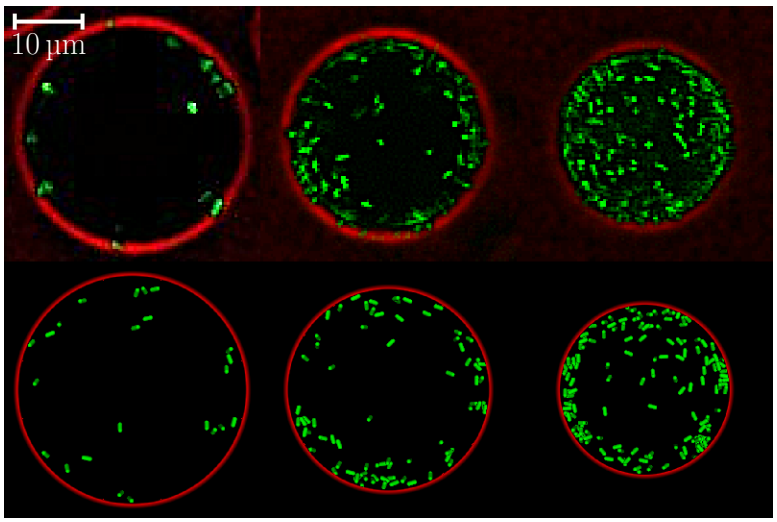


Figure 4.2 *Snapshots of* 2 µm*-thick cross-sections of droplets,* $(\phi, R) = (0.47\,\%, 17.4\,\mu\text{m})$, $(2.7\,\%, 16.1\,\mu\text{m})$, $(6.2\,\%, 14.3\,\mu\text{m})$ *(left to right). Top: experimental images with red droplet surfaces, and green bacterial cell bodies. Bottom: simulation images with the same colour assignment.*

bacterium, which was assumed to be $0.7\,\mu\text{m}^3$, consistent with typical literature values [130]. R and ϕ are considered in this analysis as the two independent variables, used to explain the observed bacterial behaviour.

4.3 Analysis of experimental data

The bacteria's positions were binned into concentric shells and counted, to form a bacterial radial number density $\rho(r)$, where r is the distance from the droplet centre. In order to allow for comparison between different droplets, the radial number density was normalised by the average number density of the droplet, ρ_0, and the radial distance normalised by the droplet size, R. In all cases a bin size of 0.7 µm was used for both experimental and simulation data.

This was done for each measured data-set of the ten taken for each droplet, and the average $\rho(r)$ taken over all data-sets. Uncertainties were estimated from the standard deviation over these data-sets. Figure 4.3a shows several distributions obtained from this method, for a roughly constant droplet radius of $R \simeq 16\,\mu\text{m}$. It can be seen that there is a peak at the droplet's oil-water interface. As ϕ increases, three changes can be seen in the peak: its excess bacterial number density decreases; its width increases; and its centre moves inwards. Meanwhile,

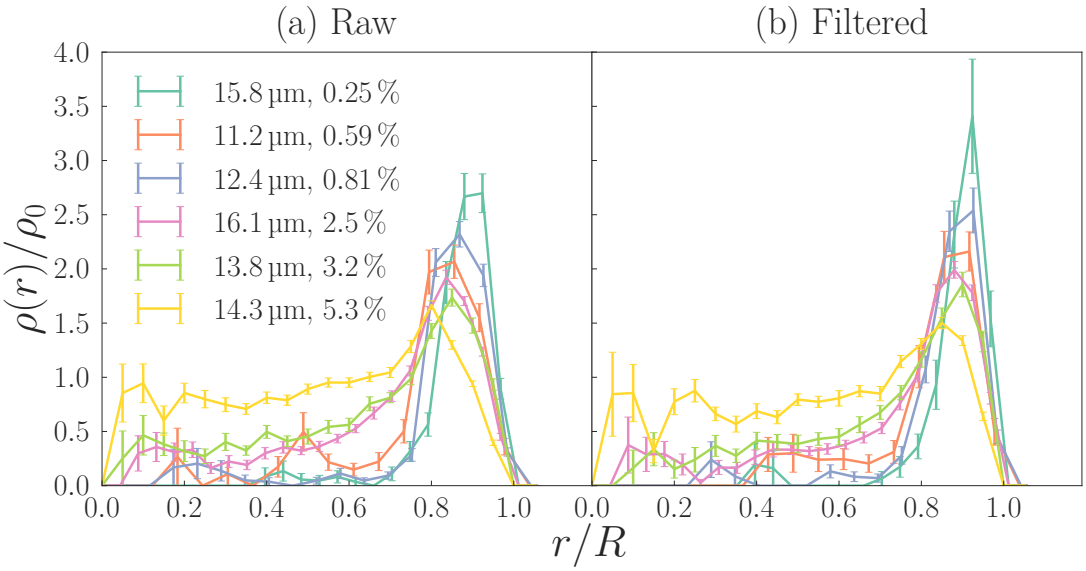


Figure 4.3 *Radial number density functions for experimental data, over an approximately constant droplet size $R \simeq 16\,\mu\text{m}$, and a range of bacterial volume fractions. (a) Data before filtering to ignore bacteria whose position was distorted by refraction through the curved droplet interface. (b) The same data after such filtering. The most obvious effect is that the inwards shift of the density peak as ϕ increases is largely removed, suggesting it is largely a measurement artifact.*

the number density in the remainder of the droplet increases uniformly.

4.3.1 Optical distortions

The imaging technique used to record the bacterial positions was found to be subject to optical distortions, caused by refraction of light crossing the oil-water droplet interface. The magnitude of these distortions were investigated by J. Arlt using ray tracing software to simulate the experimental arrangement, the results of which are shown for $R = 20\,\mu\text{m}$ in Fig. 4.4. This shows that points at large angles from the imaging axis are observed in significantly different positions than their true positions.

In order to correct for these errors, the method outlined to generate $\rho(r)$ was modified to ignore points which lie above some threshold angle, θ_{max}, from the imaging axis. This excludes points whose positions are most strongly distorted by refraction. In practice, we exclude points where $\arccos(|z|/r) > \theta_{\text{max}}$, where z is the distance from the droplet centre along the imaging axis. Figure 4.5 shows

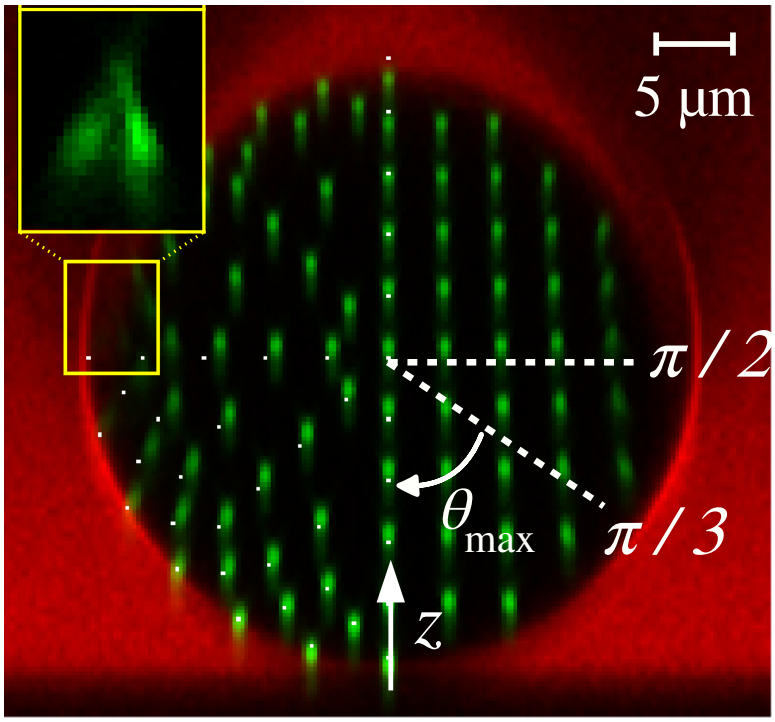


Figure 4.4 *A simulation of optical distortions due to imaging inside a spherical droplet of higher refractive index than the surrounding medium. Imaging would be done from the bottom of the image, along the indicated z-axis. When hypothetical points inside the droplet (white points in the lower-left quadrant) are imaged, their positions are displaced due to these optical distortions (green smeared points near the white points). It can be seen that points further from the imaging axis undergo a systematic inwards shift. For this reason, we filter out such points. This does not introduce bias into the results, as the system is entirely spherically symmetric; the apparent asymmetry is only introduced by the imaging equipment.*

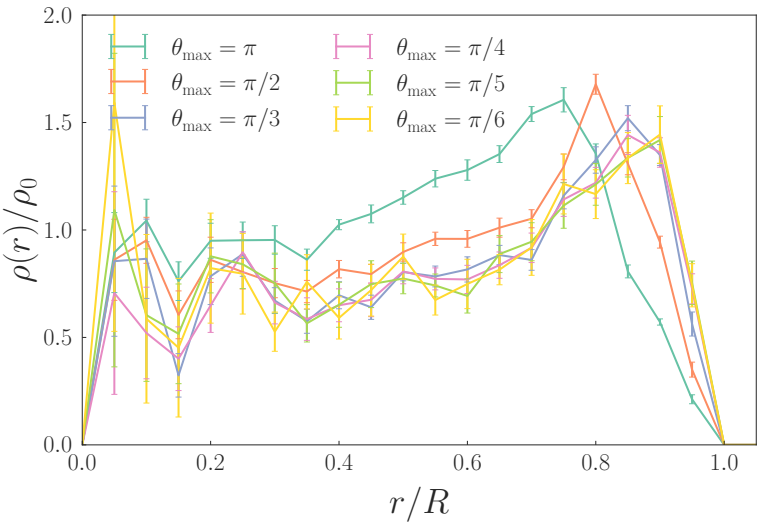


Figure 4.5 *The radial number density distribution for the same droplet, as analysed by ignoring points beyond a given angle, $\theta_{\max}$ from the imaging axis. $\theta_{\max} = \pi$ implies no filtering, while as $\theta_{\max} \to 0$, more points are excluded, with correspondingly larger error bars which can be seen. Given that the system is internally spherically symmetric, the shift seen as $\theta_{\max}$ decreases can be ascribed to the accountance for systematic measurement error. Beyond $\theta_{\max} = \pi/3$ there is minimal change in the distribution's shape, so to minimise random errors this value was chosen as an approximately optimal compromise.*

this analysis as carried out for the same droplet, using a range of values for $\theta_{\max}$. It can be seen that as $\theta_{\max}$ decreases (as we ignore more points), the location of the density peak moves outwards, towards the edge of the droplet. Since the true distribution of bacteria in the droplet is not affected by the direction from which it is imaged, without the effect of optical distortions there should be no change to $\rho(r)$ from changing $\theta_{\max}$, other than reducing the number of data-points. Therefore it was concluded that this outward movement of the peak represented a recovery of the 'true' density distribution. A negative effect of ignoring data-points is that this increases measurement uncertainties. As a compromise between minimising systematic errors from optical aberrations, and random errors from sampling a finite number of data-points, a value of $\theta_{\max} = \pi/3$ was chosen throughout the subsequent analysis.

Figure 4.3b shows the effect of this optical filtering on the experimental results: the inwards movement of the peak is reduced in magnitude, suggesting it is an artefact of the experimental arrangement.

4.4 Analytic model

We now turn to a brief description of a simple analytic model derived by A. N. Morozov to describe the observed results in this system. The model considers a bacterial solution in a cuboid of length R along one axis, and much smaller lengths along the other axes. The boundaries along the short axes are periodic. One end of the long axis represents the droplet surface; the other, the droplet centre. This arrangement approximates a narrow radial cone inside a droplet. The cuboid contains N_b uniformly distributed bacteria. Bacteria swim with a constant speed v along the long axis. The number of bacteria hitting the surface in time $\mathrm{d}t$ is therefore $N_b v\,\mathrm{d}t/R$. We assume that a bacterium hitting an unoccupied area at the surface stays there. This is motivated by the experimental observation that *E. coli* align parallel to surfaces with which they collide, as outlined in Section 2.3, which would lead to a sticking effect with the boundary conditions we consider. The probability of staying at the surface after hitting it is assumed to be equal to the fraction of the surface which is free from bacteria: $1 - N_s A_b/A$, where N_s is the number of bacteria at the surface, A_b is the surface area covered by one bacterium, and A is the total area of droplet surface. Therefore, the total number of bacteria arriving and staying at the surface during time $\mathrm{d}t$ is given by,

$$\mathrm{d}N_s = \left(1 - \frac{N_s A_b}{A}\right) N_b \frac{v\,\mathrm{d}t}{R}\,. \tag{4.1}$$

The model assumes that there are two mechanisms for bacteria leaving the surface: firstly, each bacterium may leave through its own reorientation, with characteristic time τ, that is, with rate $\gamma N_s\,\mathrm{d}t$ where $\gamma = \tau^{-1}$. Secondly, a bacterium may encounter another in a two-body 'scattering' event, which can be modelled by $\beta N_s^2\,\mathrm{d}t$. The change in surface bacteria number in time $\mathrm{d}t$ is therefore:

$$\mathrm{d}N_s = \left(1 - \frac{N_s A_b}{A}\right) N_b \frac{v\,\mathrm{d}t}{R} - \gamma N_s\,\mathrm{d}t - \beta N_s^2\,\mathrm{d}t\,, \tag{4.2}$$

where β is the scattering frequency, and γ the self-scattering frequency.

With regards to inter-bacterial scattering, consider bacteria swimming at the surface and select one bacterium as a 'target'. The probability of another bacterium hitting this target from a distance λ is $2a/2\pi\lambda$, where a is the radius of a bacterium, (which is assumed to have a circular projection on the surface). The number of

bacteria hitting the target from a thin shell $(\lambda, \lambda + \mathrm{d}\lambda)$ is,

$$\frac{2a}{2\pi\lambda}\left[\pi\left(\lambda+\mathrm{d}\lambda\right)^2-\pi\lambda^2\right]\frac{N_s}{A}=2a\frac{N_s}{A}\,\mathrm{d}\lambda\,. \tag{4.3}$$

The target can be hit only from a circle of radius $v\,\mathrm{d}t$, so that the total number of scattering events between all bacteria and a selected target is

$$\int_0^{v\,\mathrm{d}t} 2a\frac{N_s}{A}\,\mathrm{d}\lambda = 2av\,\mathrm{d}t\frac{N_s}{A}\,. \tag{4.4}$$

Since the same argument is valid for every bacterium on the surface, the total number of scattering events is proportional to

$$\frac{1}{2}N_s 2av\,\mathrm{d}t\frac{N_s}{A}=av\,\mathrm{d}t\frac{N_s^2}{A}\,, \tag{4.5}$$

where the factor of $1/2$ is introduced to account properly for the number of bacterial pairs. By comparing this expression with Eq. (4.2), we identify

$$\beta = k\frac{av}{A}\,, \tag{4.6}$$

where k is a probability that one of the bacteria participating in a scattering event would come off the surface.

Using Eq. (4.2) with the approximation Eq. (4.6), we obtain in the steady-state

$$\left(1-\frac{N_s A_b}{A}\right)N_b\frac{v}{R}-\gamma\frac{R}{v}N_s-k\frac{av}{A}N_s^2=0\,. \tag{4.7}$$

Finally, observing that $N_s + N_b = N$, where N is the total number of bacteria in the system, and introducing $\eta = N_s A_b/A$ and $\eta_0 = N A_b/A$, we obtain

$$(1-\eta)(\eta_0-\eta)=c\eta+b\eta^2\,, \tag{4.8}$$

where $b = kaR/A_b \simeq kR/a$ and $c = R/v\tau$. Note that η_0 can be significantly larger than 1.

Note that this model assumes that three-body collisions do not occur, and that a successful collision instantly dislodges a bacterium from the surface, with no possibility for secondary collisions once the bacterium is scattered away from the peak. For these reasons, the model is more valid at lower surface fractions.

4.5 Fit to the analytic model

4.5.1 Defining the peak

The analytic model described above delineates the bacteria into two populations: the 'bulk' and the 'peak'. In order to compare its predictions with the experimental data, there must be a way to perform this separation. Assuming spherical symmetry, there is effectively only one quantity which characterises each bacterium: its radial distance from the droplet centre. Therefore it is a condition on this which must be formulated to categorise it into either population.

The general approach we took was to define the peak as consisting of those bacteria with a radial distance greater than some 'peak radius', R_p. This is defined as the smallest distance at which $\rho(r)$ crosses some threshold value, ρ_b, with positive gradient. This precise definition allows the peak location to be carried out by a computer, while ignoring fluctuations at the outer edge of the droplet.

The choice of definition of ρ_b is somewhat arbitrary, and so the analysis was done with two definitions of ρ_b, to verify the robustness of the conclusions. Unless indicated otherwise, the one used hereafter is that $\rho_b = \rho_0$, that is, the peak begins where the number density first becomes larger than the average density of the entire droplet. We later show that an alternative definition does not alter our conclusions. Figure 4.6 shows the location of the peak as defined in this way for some typical droplets.

4.5.2 Cross-sectional area of a bacterium

In order to compare with the analytic model, we also must define the cross-sectional area of a single bacterium, so that we can calculate the actual and maximum fractions of the droplet surface occupied by the bacteria, η and η_0. In estimating this, we have assumed that an *E. coli* bacterium is well-approximated by the

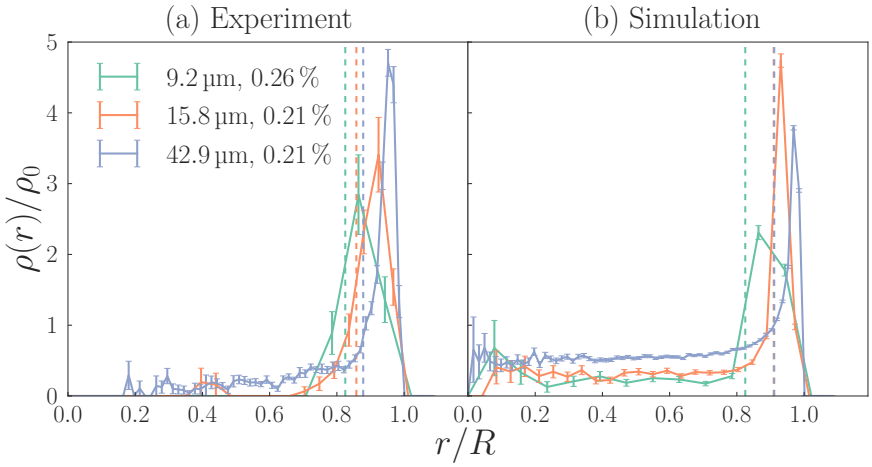


Figure 4.6 *Radial bacterial number density distributions, $\rho(r)$, normalized by the average number density of the whole droplet, ρ_0, against the radial distance from the center of the droplet, r, normalized by the droplet radius, R. As ϕ is small, the distributions can be assumed to be similar to the single-bacterium, non-interacting case. For all R shown, which span the range of that experimentally measured, the agreement between experiment and simulation is good in terms of the peak's location, width and height, however in all cases the tail of the peak on its inner edge is longer. Dashed lines indicate estimates of the peak radius R_p using the method outlined in the main text.*

geometric object known variously as a 'spherocylinder', 'spherorod' or 'capsule': a cylinder of length l and radius a, capped at both planar surfaces by hemispheres, also of radius a. The dimensions chosen for this object are $l = 1.23\,\mu\text{m}$, $a = 0.36\,\mu\text{m}$ (the justification for this choice will be outlined in Section 4.7). The maximum cross-section of this object is when viewed perpendicular to the cylinder axis; its minimum when viewed along the axis. Hence the cross-sectional area is bounded as,

$$\pi a^2 \leq A_b \leq \pi a^2 + 2al\,. \tag{4.9}$$

Since much experimental work suggests that a swimming *E. coli* bacterium aligns parallel to a surface, we have chosen to assume each bacterium occupies the upper bound on the above limits. For these values this implies an occupied area of $1.3\,\mu\text{m}^2$.

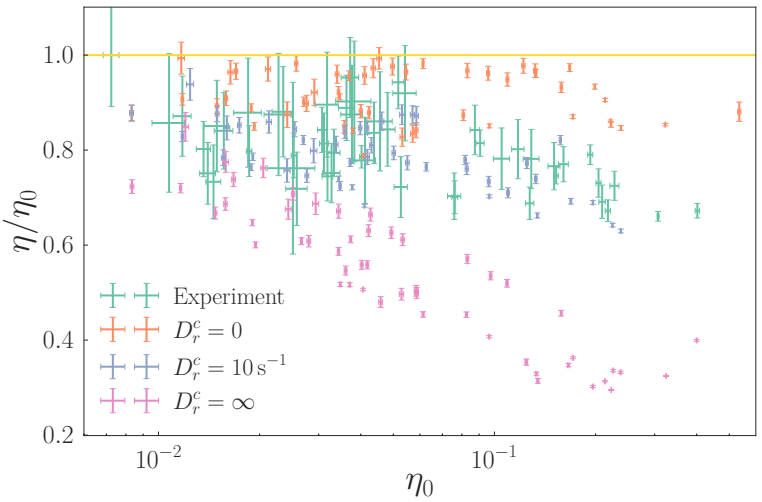


Figure 4.7 *The fraction of bacteria within the peak, η_0/η, as a function of the 'surface area filling fraction', η_0. This is the fraction of the droplet's surface area which could be covered by all bacteria in the droplet. The quantity can be seen as analogous to the* volume *filling fraction, ϕ. For the experimental data (green), it can be seen that as the bacteria's ability to occupy the droplet surface increases, there is a decrease in the fraction that in fact do occupy it. Also shown is simulation data in the strong (pink) and weak (orange) scattering limits, and for $\mathrm{D}_r^c = 10\,\mathrm{s}^{-1}$, the optimal fit to the experimental data (purple).*

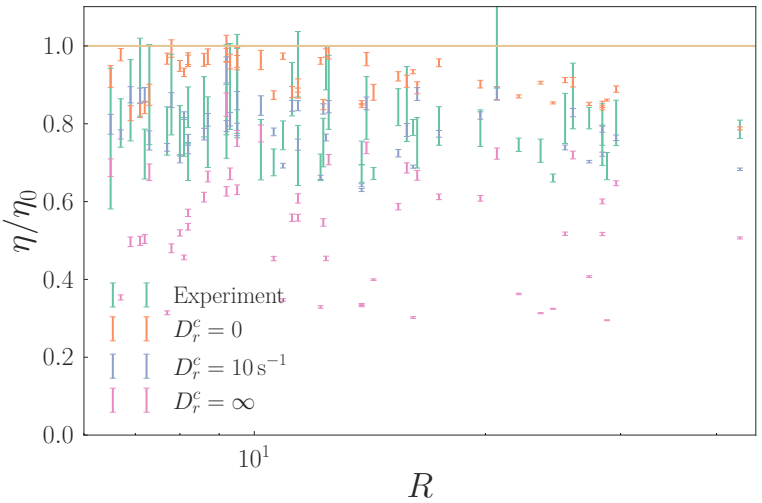


Figure 4.8 *The fraction of bacteria within the peak, η_0/η, as a function of the droplet radius, R. Data displayed as in Fig. 4.7. The analytical model predicts that the peak fraction should decrease with increasing R, but there is no clear dependence in any data-set shown. This is to be expected, however, as the point at which the decrease becomes significant is when the persistence length of the bacterium becomes shorter than the droplet size. For $\mathrm{D}_r \simeq 0.062\,\mathrm{s}^{-1}$, the persistence length is $l_r \simeq 200\,\mu\mathrm{m} \gg R$ for all R considered here.*

4.5.3 Fitting the model to the data

Figures 4.7 and 4.8 show the dependence of η/η_0 (which is exactly equivalent to the fraction of bacteria in the peak) on the two independent variables under study: η_0 (representing the average density of bacteria in the droplet) and R. From Fig. 4.7 it can be seen that increasing bacterial number density causes a smaller fraction of them to be found in the peak. This is in line with the qualitative results seen in both Fig. 4.2 and Fig. 4.3. Figure 4.8 also indicates the same decreasing trend for increasing R, however this is less clear. Given that the persistence length of smooth-swimming *E. coli* is larger than even the largest droplets used in the experiments, the effect of increasing droplet size is expected to be weak.

This data was then fitted to Eq. (4.8) by non-linear least squares analysis. The least squares approach in general attempts to find values for the parameters in a given model, such that the sum of the squared differences between the model's prediction, and the actual data, is minimised. To be more concrete: in this particular case, the approach aims to minimise,

$$S = \sum_{i}^{m} \left(\eta_i - \eta(\eta_{0,i}; k, \tau)\right)^2 , \tag{4.10}$$

where η_i is the ith experimentally measured value, and $\eta(\eta_{0,i}; k, \tau)$ is the model's predicted value with a given choice of values for the parameters k and τ.

In the model considered here, the dependent variable is a non-linear function of both parameters. As such, so-called non-linear least squares fitting is required. This fitting was done using the Levenberg-Marquardt algorithm as implemented in the 'MINPACK' library. This fitting gave the following parameter estimations: $k = 0.08(2)$, $\tau^{-1} = 0.14(3)\,\mathrm{s}^{-1}$. The values of η predicted by the analytic model for the parameter sets sampled by the experiments is shown in Fig. 4.9, indicating a decent fit, particularly at large η_0.

4.5.4 Robustness of results to changes in peak definition

Since the definition of the 'peak radius' is somewhat arbitrary, analysis of both the simulation and experimental data was carried out again under a different definition, to verify that the broad conclusions are unchanged. Here, instead of

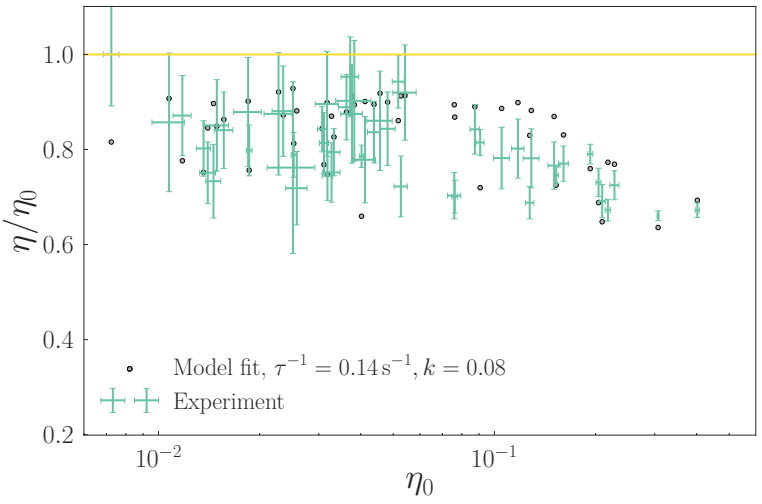


Figure 4.9 *Experimental data as in Fig. 4.8, shown against the values predicted for each data-point by the analytic theory, based on parameters fitted to the experimental data.*

defining $\rho_b = \rho_0$, we choose,

$$\rho_b = \bar{\rho}(r) + \kappa \left(\rho_{\max} - \bar{\rho}(r)\right) \ , \tag{4.11}$$

where $\bar{\rho}(r)$ is the median density over the radial density function, $\rho_{\max}$ is its maximum value, and κ is a parameter between 0 and 1 which determines how close to the maximum the distribution is required to be, for the peak to be considered to have begun. Note that the median is not weighted by the shell volume represented by each $\rho(r)$: $\rho(r)$ is considered simply as a 'flat' sequence of values.

The choice of κ was made by taking the mean of two independent judgements of R_p, done by visual inspection of $\rho(r)$ for all experimental data-sets, and choosing the κ which best agreed with these estimates; this resulted in $\kappa = 0.2$. For this value, Fig. 4.10 shows the estimates for the peak occupancy fraction, as compared to the alternative peak definition method. This gives an estimate for the model parameters of $k = 0.07(4)$, $\tau^{-1} = 0.15(4)\,\mathrm{s}^{-1}$. These results were judged to be sufficiently similar to the results using the previous peak definition to demonstrate robustness of the analysis.

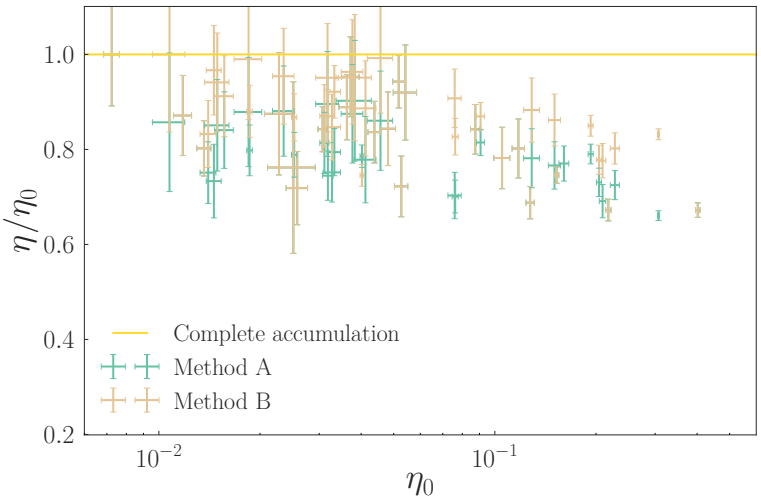


Figure 4.10 *The peak fraction as a function of surface coverage, comparing two different methods of determination for the lower radial distance bound on the 'bacterial peak'. Method A (green) defines the peak as where the radial density intersects the average density. Method B (brown) defines the peak where the density intersects ρ_b, as defined in Eq.* (4.11).

4.6 Description of simulation model

The analytic model gives no description of how the bacteria are distributed in the droplet. To this end, we construct a microscopic, agent-based model of the system.

The water droplet is represented as an impenetrable, thin spherical shell, which is at rest. Each bacterium is approximated as an impenetrable spherocylinder. Each spherocylinder is polar along its long axis: one end is considered to be the bacterium's 'head'; the other its 'tail'. The flagella are not considered to occupy space.

An advantage of using spherocylinders in simulations is that it makes implementation simple and fast. To determine if two spherocylinders intersect, the problem is equivalent to finding the shortest distance between two finite line segments. This is a simple algorithm, and this distance can then be compared to the radius of the spherocylinders to check for intersection [131].

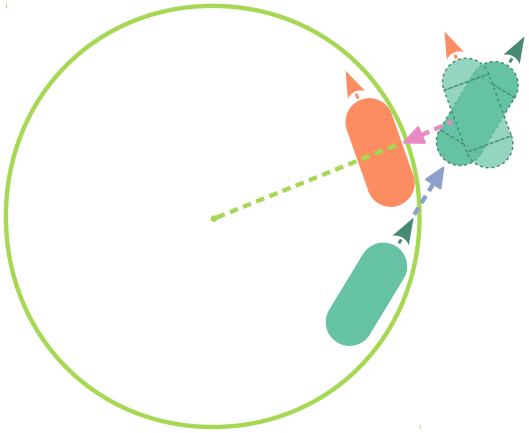


Figure 4.11 *The algorithm used to implement velocity alignment of bacteria with the droplet surface. At each step, an agent's position is displaced without considering the presence of the droplet. If any part of the agent then lies outside of the droplet interface, the velocity component perpendicular to the interface at the agent's centre is set to zero, and the velocity scaled to maintain its previous speed. The agent's radial position relative to the droplet centre is then reduced such that it is touching the droplet interface at each pole.*

4.6.1 Interaction of a bacterium with the inner edge of a water droplet

It has been observed in experiments that motile *E. coli* orient parallel to nearby surfaces, and swim along them. It is this behaviour which results in their non-uniform steady-state distributions in many environments. As such it is important to include this feature in the simulations.

The droplet surface is modelled as impenetrable. When, in a given time-step, a bacterium's displacement would cause it to intersect the boundary of the droplet, the spherocylinder is rotated such that its orientation vector is parallel with the surface at its point of intersection. That is, the component of its orientation vector which is normal to the droplet surface is set to zero, and the remaining vector components scaled to maintain unit magnitude. The radial component of the bacterium's position is then scaled such that it is just touching the droplet boundary (Fig. 4.11).

The reason for using this updating rule to characterise the bacterium's interaction with the droplet surface is that there are both torque and force components to the interaction. To match experimental findings, the bacterium coming into contact with the droplet surface must cause it to both align with the surface, and be

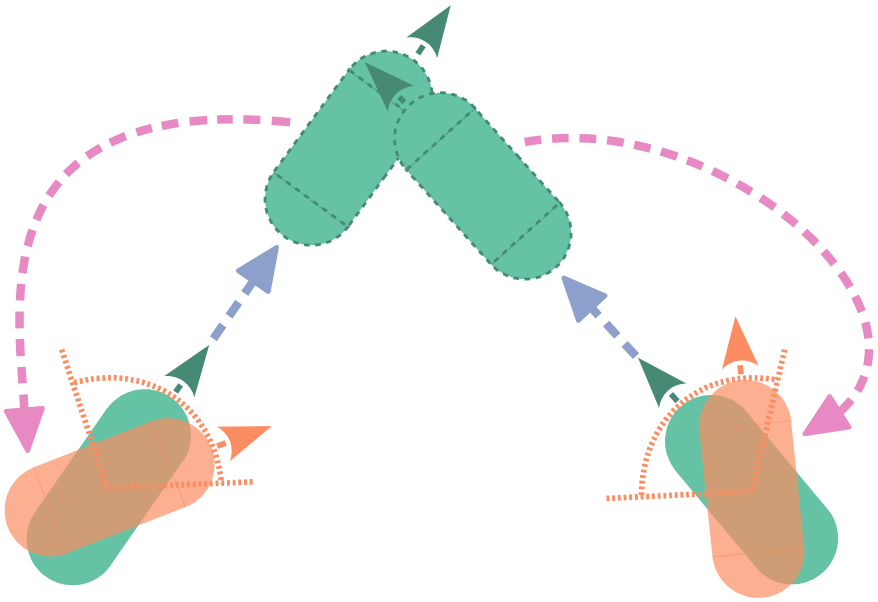


Figure 4.12 *The algorithm used to implement collisions between bacteria. At each step, all agents' positions are displaced independently, without considering their interactions (purple). If a pair of agents intersect, their positions are reset to their original positions (pink), and they are rotated about that point by an angle chosen from a normal distribution of a width determined by* D_r^c *(orange).*

repelled from it. Although it would be possible to define a polar interaction which specifies the torque and force between the bacterium and droplet surface, as a function of the angle between the swimming axis and surface normal, and the distance between bacterium and surface, this would introduce a number of choices of model and parameters, which are not well constrained by the available experimental data. As such, we choose the above idealisation which reduces the number of free parameters we must take into consideration.

4.6.2 Interactions of bacteria with each other

With regards to inter-bacterial collisions, the bacteria were again approximated as impenetrable spherocylinders. This means that if the displacement of an agent due to swimming, translational or rotational diffusion leads to an intersection occurring between particles, that displacement of position or direction does not occur.

In addition to this hard-core interaction, an additional interaction is imposed in order to implement scattering between bacteria. If the dynamics of a bacterium

would cause it to intersect with another bacterium (that is, if there is a collision between them), an additional contribution to its rotational diffusivity is applied, D_r^c (Fig. 4.12). In practice this interaction is not expected to in fact be 'diffusive': a small number of large-angle 'deflections' is likely to be a more realistic scenario. However, modelling the interaction as a diffusive process allows the implementation to be independent of the simulation's time-step, as it causes a deflection whose magnitude scales with the time-step.

4.6.3 Initial conditions

The system was initialised with the bacterial positions uniformly distributed in the droplet. Their orientations were also chosen uniformly. This configuration was arrived at using an implementation of the Metropolis-Hastings algorithm, as outlined in Section 3.5, applied to pack the required number of spherocylinders inside a sphere.

4.7 Choice of simulation parameter values

It was determined that the bacteria used in the experiment had cell volumes of approximately $0.7\,\mu\mathrm{m}^3$, in approximate agreement with literature values [132]. By measurement from microscope images such as Fig. 4.2, an aspect ratio of 2.7 was chosen. Given these two constraints, the spherocylinder dimensions are required to be, $l = 1.23\,\mu\mathrm{m}$, $a = 0.36\,\mu\mathrm{m}$.

Values for the diffusive characteristics were taken from literature values for *E. coli*: $\mathrm{D} = 0.2\,\mu\mathrm{m}^2\,\mathrm{s}^{-1}$, and $\mathrm{D}_r = 0.062\,\mathrm{rad}^2\,\mathrm{s}^{-1}$ [133].

Differential dynamic microscopy (DDM) was performed to measure the swimming speed distribution of the bacteria within individual droplets and the fraction of non-motile organisms. An average swimming speed inside a droplet was found to be $\bar{v} \simeq 13.5(7)\,\mu\mathrm{m\,s}^{-1}$ at cell concentrations similar to those used in the experiment. Hence a value of $v = 13.5\,\mu\mathrm{m\,s}^{-1}$ was used in the simulations.

All simulations were run with a time-step of $\Delta t = 1\,\mathrm{ms}$. This means a bacterium travels its own length in $n_s = (l + 2a)\,/\,(v\Delta t) \simeq 150$ time-steps; the bacterium randomises its direction through rotational diffusion in $n_\theta \simeq 1/\,(\mathrm{D}_r\Delta t) \sim 10^4$ time-steps, and the bacterium diffuses its own radius in $n_D \simeq a^2/\,(6D\Delta t) \simeq 300$

time-steps. These numbers were considered sufficiently large to justify the choice.

The simulations were run for $t = 500\,\mathrm{s}$, which was sufficient to reach a steady-state for all ϕ and R investigated. Bacterial positions were recorded for the last 100 s, every 1 s, and the radial density distributions shown reflect the average over these recordings; their errors are taken from the standard deviation over them. These errors may be slightly under-estimated since the data-sets are not strictly independent, however the errors are sufficiently small that this was not considered to be an issue.

4.8 Radial density distributions

At low volume fractions, inter-bacterial collisions occur only rarely, and the results become independent of simulation parameters such as the bacteria's aspect ratio (though their volume remains important as it significantly affects their diffusive characteristics), and D_r^c. Figure 4.6 shows a comparison between experiment and simulation at these low volume fractions. The qualitative features of the experiment are reproduced by the simulations: a peak at the edge of the droplet around 3 µm wide, with a number density around three times higher than the average.

We next turn to higher volume fractions, where inter-bacterial collisions are more important. Figure 4.13a shows how the peak evolves as the volume fraction is increased, at constant $R = 16\,\mu\mathrm{m}$, with no collisional scattering, $\mathrm{D}_r^c = 0$. This means that at a collision between bacteria, they 'stall' until bulk translational and rotational diffusion frees them to continue on their path. It can be seen that increasing the number of bacteria has almost no effect on their distribution, in contrast to the experimental results (Fig. 4.3b). This implies that the peak is decreased at these volume fractions not because of exclusion effects – the interface simply 'filling up' with bacteria and there being no more room available – but instead because of explicit interactions between them.

Figure 4.13b shows results from the same set of simulations, but for the case where $\mathrm{D}_r^c = \infty$. This means that any collisions between bacteria cause them to completely randomise their direction. It can be seen that there is a strong dependence of the peak on ϕ, leading to an approximately uniform distribution at $\phi \simeq 4\,\%$.

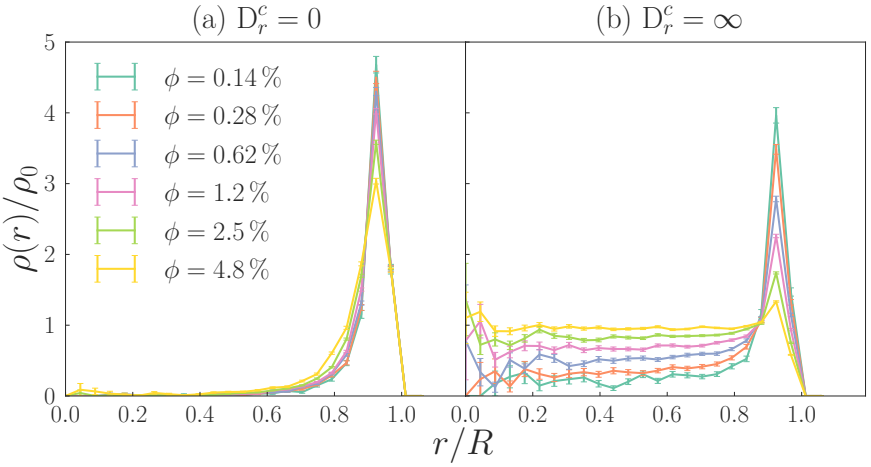


Figure 4.13 *Radial bacterial number density distributions from simulations with a constant droplet size $R = 16\,\mu\text{m}$, and variable volume fraction, ϕ. (a) Results with no inter-bacterial collisional rotational noise. The small reduction in the size of the peak with increasing ϕ is due to exclusion of incoming bacteria from the surface by those already there. (b) Results for the upper limiting case of inter-bacteria collisional rotational noise, where bacteria completely randomise their direction on colliding with one another. The much larger reduction in the size of the peak is due to bacteria already at the surface meeting one another, and being scattered away into the bulk.*

4.9 Statistical moments

Additional measures of the data which aid comparison are the statistical moments of the bacterial radial positions. The first and second statistical moments – the mean and the variance – of the radial distance of the bacteria from the centre of the droplet, as a function of the volume fraction, are shown below.

It is useful to compare the observed mean and variance with that expected if the bacteria were distributed uniformly. If this were the case we should expect a probability distribution, $p(r) \propto r^2$ inside the droplet. This implies a probability density,

$$p(r) = \begin{cases} \dfrac{3}{R^3} r^2, & r < R \\ 0, & r \geq R \end{cases} . \tag{4.12}$$

We can now find the uniform expected mean according to its definition,

$$\begin{aligned}\bar{r} &= \int_0^\infty p(r) r \, \mathrm{d}r \\ &= \frac{3}{R^3} \int_0^\infty r^3 \, \mathrm{d}r \,, \\ &= \frac{3}{4} R\end{aligned} \tag{4.13}$$

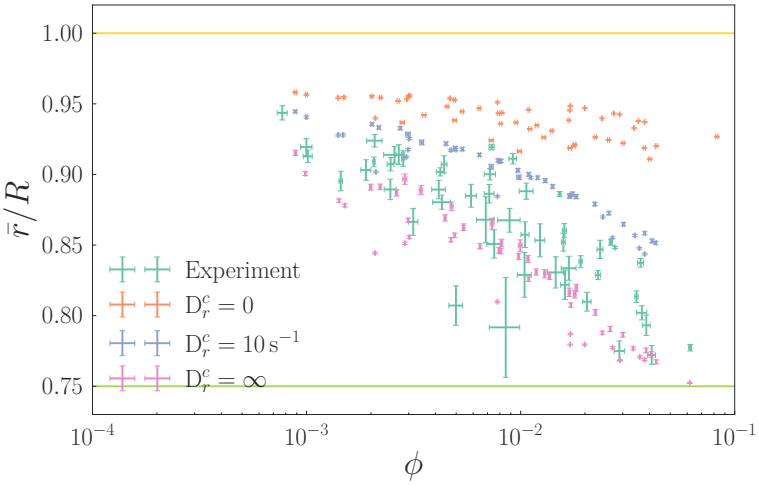
and using this result we can find the uniform expected variance,

$$\begin{aligned}\mathrm{Var}(r) &= \int_0^\infty p(r) r^2 \, \mathrm{d}r - \bar{r}^2 \\ &= \frac{3}{R^3} \int_0^\infty r^4 \, \mathrm{d}r - \left(\frac{3}{4} R\right)^2 . \\ &= \frac{3}{80} R^2\end{aligned} \tag{4.14}$$

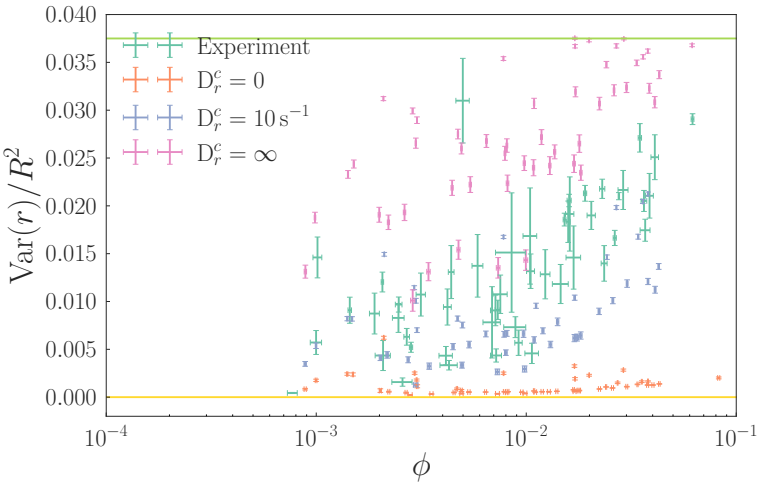
Figure 4.14a and Fig. 4.14b show the statistical moments as calculated for experiment and simulation. Both results imply a larger scattering parameter than the one fitted to Fig. 4.7, but the upper scattering limit ($\mathrm{D}_r^c = \infty$) is also inconsistent with the experimental data. We conclude from this that the fitted parameter based on the analytic model, $\mathrm{D}_r^c = 10\,\mathrm{s}^{-1}$, should be taken as a lower bound on the strength of inter-bacterial scattering, however the fact that the upper scattering limit is also inconsistent with the experimental data supports the conclusion that scattering is not entirely randomising.

4.10 Direct measurement of the scattering probability

In order to test the validity of the scattering probability determined by fitting to the analytic model, additional inspection of the simulation results was done in an attempt to determine the probability of an inter-bacterial collision leading to bacteria leaving the peak, through a more microscopic analysis.



(a)



(b)

Figure 4.14 *Statistical moments of the bacterial radial positions as a function of the volume fraction ϕ. The expected values for a uniform distribution (green), and for complete accumulation (yellow) are indicated. (a) Mean radial position, normalised by the droplet radius, R. (b) Radial position variance, normalised by the droplet radius. Both measures show good agreement between fitted model simulations and experiment. The mean approaches its uniform value, as does the variance to a lesser extent, bolstering the conclusion that the accumulation is largely lost when $\phi \simeq 4\,\%$.*

4.10.1 Methodology

During a simulation run, on the occurrence of a collision at time t_i, a 'countdown timer' was attached to each colliding bacterium, set to a relaxation time, τ, and their current radial distance recorded, R_i. After the time τ had elapsed, the new radial distance was recorded, R_f. If additional collisions occur during $[t_i, t_i + \tau]$, the timer was *not* reset.

To convert this into a peak scattering parameter, we must define a radial distance at which the peak begins, R_p. Only initial collisions occurring within the peak were considered, $R_i > R_p$. The collision events were then separated into those where the particles finished in the peak once again, $R_f > R_p$, and those where the particles finished in the bulk, $R_f < R_p$. The numbers of each event, n_{pp} (peak-to-peak) and n_{pb} (peak-to-bulk) were counted, and from this the fraction of peak collisions resulting in scattering into the bulk could be calculated,

$$\tilde{p} = \tilde{k} + \tilde{f} = \frac{n_{pp}}{n_{pp} + n_{pb}} \,. \tag{4.15}$$

This equation separates the probability of scattering into the bulk, $\tilde{p}$, into that from inter-bacterial scattering, $\tilde{k}$, and that from scattering through diffusion, $\tilde{f}$. $\tilde{f}$ can be measured by running the same simulation for a single bacterium, so that there are no collisions, $\tilde{k} = 0$, and measuring the probability that it moves from the peak to the bulk in any given period τ. $\tilde{k}$ can be associated with the scattering parameter, k, in the analytic model.

If τ is very short, the bacterium does not have time to leave the peak after a collision, even if scattered directly into the bulk. If τ is very long, the effect of the collision becomes swamped by diffusion, and/or the bacterium leaves and re-enters the peak.

As a compromise between these extremes, we chose τ to be the ballistic crossing time of a bacterium across the droplet's radius. For example, for a droplet with radius 16 µm, with a swimming speed of $13.5\,\mu\mathrm{m\,s^{-1}}$, $\tau = t_{\mathrm{cross}} \simeq 1.2\,\mathrm{s}$. This reduces the likelihood of bacteria re-entering the peak, and is long enough that a bacterium has time to leave the peak if it has been scattered away by the collision.

Simulations were run, with this analysis done during the simulation, for a droplet with $R = 16\,\mu\mathrm{m}$ (as this is approximately the most common size of droplet used

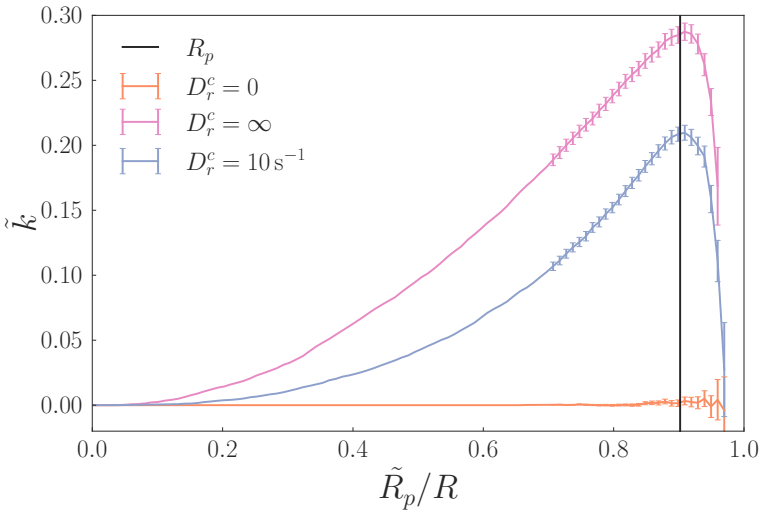


Figure 4.15 *The directly measured probability of a collision between bacteria in the peak leading to one of them entering the bulk, as a function of where the peak is defined. For the infinite-scattering data, $\tilde{k}$ has a maximum close to the value defined by Section 4.5.1, which serves to validate the approach. As expected, the zero-scattering data is consistent with $\tilde{k} = 0$, regardless of where the peak is chosen to lie.*

in the experiment), and two bacteria to ensure no three-body collisions would occur. The simulation was run for a sufficiently long time for enough collisions to occur on which to do good statistics, and then $\tilde{p}$ was calculated. The same simulation was then run with a single bacterium. The difference between the two-body and one-body cases represents the probability of scattering away from the surface solely due to inter-bacterial collisions, $\tilde{k}$.

Figure 4.15 shows this quantity as a function of the peak cut-off $\tilde{R}_p$. The natural value to use in the analysis is $\tilde{R}_p = R_p$ as defined in Section 4.5.1, which is indicated. For this value, for the infinite scattering case, $\tilde{k} = 0.28(2)$. For the zero-scattering case, $\tilde{k} = 0.003(9)$, consistent with $\tilde{k} = 0$ as would be expected. For the fitted scattering parameter, $\mathrm{D}_r^c = 10\,\mathrm{s}^{-1}$, $\tilde{k} = 0.210(6)$. The value of k obtained by fitting the simulation model to the experimental data was $k = 0.08$. Therefore the direct measurement method suggests that the scattering probability implied by the analytic model may be an underestimate, and in fact bacteria have a higher probability of leaving the droplet surface due to collisions.

4.11 Correlation functions

In all the analysis done above, the distribution of bacteria is assumed to be spherically symmetric about the droplet centre. This is likely to be true over an ensemble of droplets, however in a particular case there may be anisotropies arising from correlations of position between bacteria: that is to say, the bacteria may spontaneously break symmetry by clustering in particular regions of the droplet. In order to investigate this, two analyses were carried out: firstly, a bacterial radial distribution function was calculated for all bacteria within the droplet; secondly, an angular distribution function was also calculated, for those bacteria at the droplet interface.

4.11.1 Distance correlations

In general, a 'radial distribution function' (also known as a 'pair correlation' function), measures the likelihood of finding two particles at a particular distance from one another, as a function of that distance. To make the distinction clearer between this function, and $\rho(r)$, we will refer to this as the 'pair correlation' function from here onwards. To be clear, the radial density distribution, $\rho(r)$, measures the probability of finding a bacterium at a particular distance from the centre of the *droplet*, while the pair correlation function (PCF) measures the probability of finding a bacterium at a particular distance from *another bacterium*, averaged over the ensemble of all bacteria.

To be more precise, the value of the PCF, $g(r)$, at a given r represents the probability of finding a particle at a distance r from another particle, normalised by the probability expected in a system of particles whose positions are completely uncorrelated. In practice, this can be calculated by making a list of scalar distances between each pair of particles, binning them to form a histogram of bin width Δr, representing the number of bacteria within a given spherical shell between $[r, r + \Delta r]$, $n(r)$, and normalising by the expected value of this function in an uncorrelated system, $n_I(r)$.

In a spherically symmetric, infinite system (an ideal gas) $n_I(r)$ is trivial to calculate: the expected number of particles in a shell of radius r and width Δr is simply the average density of the system, ρ_0, multiplied by the shell volume: $n_I(r) = 4\pi r^2 \Delta r \rho_0$. However the system under consideration is not infinite, since it

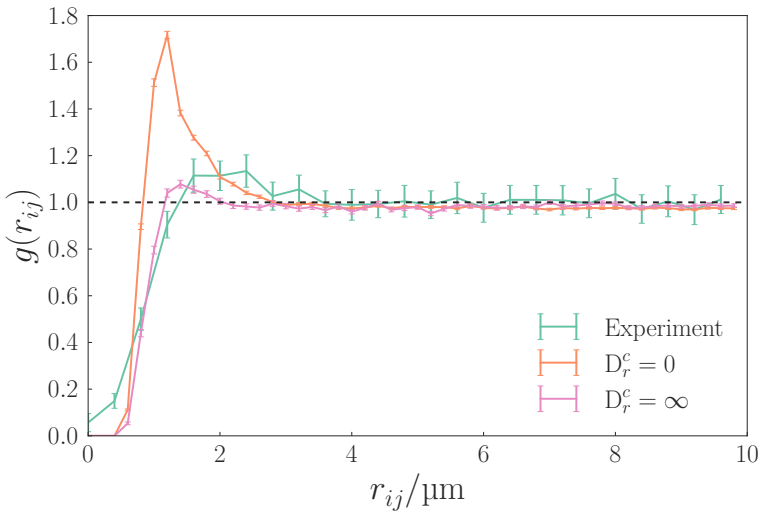


Figure 4.16 *Pair correlation function for bacterial positions within a droplet of radius $R = 28.8\,\mu\mathrm{m}$, with volume fraction $\phi = 1.7\,\%$. r_{ij} indicates that this is the pair-wise distance, to distinguish it from r, the radial distance from the droplet centre used elsewhere. Although not shown, qualitatively similar results were found for other values of R and ϕ; the particular data shown were chosen because the large number of bacteria (~ 2500 in each measurement) allows for better statistics. We use a bin size of $\Delta r = 0.2\,\mu\mathrm{m}$ for the simulation data, and $\Delta r = 0.4\,\mu\mathrm{m}$ for the experimental (as less data is available, a smaller bin size would reduce legibility).*

is known that for $r > 2R$, $n(r) = 0$, and for smaller r there are biases introduced by the finite size of the droplet. In addition, for the experimental data there is an additional loss of spherical symmetry, due to the fact that the optical correction causes all data to lie within a cone originating in the droplet centre.

For these reasons, $n_I(r)$ was calculated numerically. This was done by picking a large number ($\sim 10^4$) of points uniformly within the domain of the actual data. In order to ensure the correlation functions were comparable between simulation and experiment, *both* sets of data were made to lie within the same domain, meaning that the simulation data had the same optical correction applied. This uniform distribution of points can then be used to calculate $n_I(r)$, and used to normalise the actual data. This uniform point distribution was formed simply through rejection-sampling, where points were picked uniformly on a cube of length $2R$ (by picking three values from a uniform distribution), and those found to lie within the valid domain were retained. This method is not efficient, but was sufficient for the number of samples that were found to be necessary to give a smooth estimate of n_I (10^4), and is guaranteed to form a uniform distribution of points [134].

Figure 4.16 shows the result of this calculation for a droplet with a sufficient volume fraction for inter-bacterial collisions to be significant. As might be expected, in all cases there is strong anti-correlation at distances smaller than the typical size of the bacterium, $\sim 1\,\mu\text{m}$, due to volume exclusion. In all cases there is also positive correlation at distances from $1\,\mu\text{m}$ to $2.5\,\mu\text{m}$ for both simulations and for experiment.

For the sake of comparison, for a distribution of static hard spheres of radius σ, the peak in the correlation function at $r = \sigma$ is given [135] by,

$$g(r = \sigma) = \frac{1 - \phi/2}{(1 - \phi)^3}\,. \tag{4.16}$$

Making the rough approximation that the bacteria can be modelled as spheres, this would give a peak for the droplet considered here, where $\phi = 1.7\,\%$, of $g(\sigma) = 1.04$. This shows that for $\mathrm{D}_r^c = \infty$, where the peak in $g(r)$ is around 1.1, the correlation is only slightly more than would be expected due only to volume exclusion, without active effects.

For $\mathrm{D}_r^c = 0$, the positive correlation is much stronger, indicating that bacteria spend a significant amount of time near to each other. The length scale of the correlation is, however, similar between both sets of simulation data, $\sim 2\,\mu\text{m}$.

The experimental data indicates a better agreement with the strong-scattering model, in terms of the peak in $g(r_{ij})$ at ~ 1.15. However the length scale of the correlation is longer, $\sim 3.5\,\mu\text{m}$. This may be due to a greater degree of anisotropy in the bacteria's shape than is incorporated into the model (due to its neglect of the flagella), or due to hydrodynamic interactions (which can be significant at separations up to $15\,\mu\text{m}$ [136]).

4.11.2 Angular correlations at the interface

We next turn to correlations in the location of the bacteria within the peak density region of the droplet, near the interface. This was measured by calculating the angle between each pair of vectors, where the vectors represent the positions of the bacteria in a coordinate system whose origin is the droplet's centre. The angular

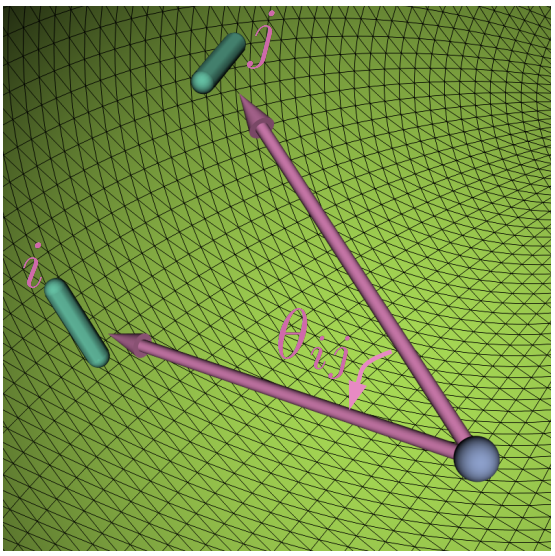


Figure 4.17 *Schematic showing the meaning of the angular separation, θ_{ij}, between two bacteria in the droplet. The angle is found between the vectors (pink) pointing from the droplet centre (purple) to the two bacteria in question (teal). The droplet surface is indicated in green.*

separation between two vectors in three dimensions can be found by computing,

$$\theta_{ij} = \arctan\left[\frac{\left\|\mathbf{r}_i \times \mathbf{r}_j\right\|}{\mathbf{r}_i \cdot \mathbf{r}_j}\right], \tag{4.17}$$

where $\mathbf{r}_{i(j)}$ is the $i(j)$-th bacterial position vector [137]. This is illustrated in Fig. 4.17.

This quantity was calculated for all bacteria where $\|\mathbf{r}_i\| > R_p$, that is, those within the peak as defined previously, and the same method described above was applied to form an angular correlation function (ACF), $g_\theta(\theta_{ij})$, shown in Fig. 4.18. It is notable that $g_\theta(0) > 0$: since the bacteria cannot intersect each other, this implies that they are at the same point in the droplet but at different radial distances. This implies that the density peak consists of more than a single 'layer' of bacteria. Nevertheless, there is still a net anti-correlation for $\theta_{ij} < 0.04\,\text{rad}$. The distance between two points on the surface of a sphere is given by $d = r\Delta\sigma$, where d is the distance along the surface, r is the radius of the sphere, and $\Delta\sigma$ is the angular separation of the points [138]. Assuming bacteria are at the interface, this gives a separation of $r_{ij} = \theta_{ij}R \simeq 1.4\,\mu\text{m}$, which is sensible in light of the bacteria's size.

Similarly to the PCF, for the simulated data with $\mathrm{D}_r^c = \infty$ there is no correlation

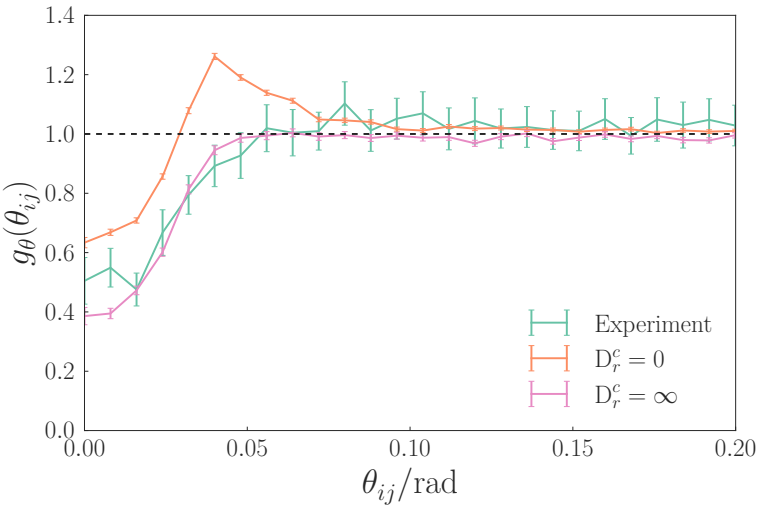


Figure 4.18 *Angular correlation function for bacterial positions within the peak density region of the droplet.*

beyond that expected from volume exclusion. For $D_r^c = 0$ there is a peak at 0.04 rad. The experimental data is consistent with an absence of clustering.

4.12 Conclusion

First, we have shown the distribution of motile, smooth-swimming *E. coli* in spherical droplets, at low bacterial densities. In such a scenario, the bacteria form a well-defined layer at the interface approximately 4 µm thick. This distribution is well-predicted by agent-based simulations assuming a kinematic interaction of bacteria with the droplet interface, free from hydrodynamics, whose results suggest that when the droplet size is smaller than the bacterial persistence length, the fraction of bacteria in the peak is determined largely by the rate of rotational diffusion of the bacteria.

Second, at increased bacterial densities, we have demonstrated the emergence of many-body interactions causing the density peak to be diminished. Simulations suggest that this is due to scattering between bacteria within the layer causing re-orientation of the swimming direction, leading to persistent 'traffic' as these scattered bacteria cross the droplet bulk. The nature of these interbacterial interactions have been probed by modelling them as a scattering-induced increase in a swimmer's rate of random re-orientation. The parameter introduced by this model was fitted to the experimental data, with the result that $D_r^c = 10\,s^{-1}$. Additional inspection of the statistical moments of the data suggests this may be

an underestimate, however complete inter-bacterial scattering is not consistent with the analysis, in contrast with many previous assumptions.

The agent-based model we have used could be extended to account for the spatial exclusion effects caused by the presence of bacterial flagella, which we have neglected in this analysis. We also assume the bacteria's diffusive characteristics are isotropic, however translational diffusion transverse to the cell body is expected to be faster than diffusion along the swimming axis. Where the bacteria are swimming parallel to the droplet surface, this would enhance the rate of diffusion away from the surface, as this direction is transverse to the cell body.

Chapter 5

Bacterial pattern formation via auto-chemotaxis in confining geometries

5.1 Introduction

In this chapter we again look at the distribution of interacting bacteria in confined environments, but where the interaction is on a longer length scale than that of explicit volume exclusion and scattering. Instead the interaction is through a mediating chemical field, gradients of which attract the bacteria.

Chemotaxis alone is not sufficient to result in significant bacterial interactions — one can study distributions of dilute suspensions of bacteria in an externally imposed chemical gradient — but here the study is of the case where this chemical is secreted by the bacteria themselves. This can lead to complex non-linear behaviour from simple systems. A positive feedback effect arises, where an upwards fluctuation in the density of bacteria leads to more of the chemoattractant being secreted at the location. This leads to a gradient which attracts more bacteria, increasing the size of the density fluctuation, and so on. The size of the initial density fluctuation is likely to be important however, due to the fact that the secreted chemical diffuses away, meaning that small spontaneous chemical gradients may be washed out before the bacteria's chemotaxis leads to a response.

There have been many previous studies of this kind of feedback behaviour, however

its effects in geometrically complex environments are less well understood, despite the fact that many bacteria live in such environments. The time-asymmetric interaction of bacteria with planar surfaces leads to increased bacterial density close to the surfaces (as seen in the previous chapter). It is supposed that the presence of surfaces may affect the spontaneous clustering outlined above in two ways: firstly, it may lead to larger density fluctuations than would arise in a uniform environment, meaning that the threshold of chemotactic strength required for clustering is lowered; secondly, it should localise the density fluctuations in particular regions of space, close to the surfaces, meaning that the *location* of the clusters is no longer completely random.

The overall aim of the investigation is to establish a minimal realistic model of the dynamics of chemotactic bacteria in a uniform, periodic system, and then introduce obstacles to see their effect on the tendency of the bacteria to form dense aggregates, as well as their density, number, coarsening and location. In the following sections, we first establish this minimal realistic model, using an agent-based approach, to explain the effect of interactions of bacteria with obstacles in their environment, and to predict where gradients in an attractant molecule will arise. we then outline the dynamics of a concentration field representing such a molecule. Results are shown for a one-dimensional version of the model, and compared with analytic results from a coarsened drift-diffusion model. These are then generalised to two dimensions, and similarities and differences are presented due to the change in dimensionality. Environmental obstacles are then placed in the two-dimensional environment, inspired by the work of Park, Austin *et al.* [66, 68], and the effect on the bacteria simulated. The simulated environments progress in complexity from a single trap, through to multiple traps, and finally to a randomly generated maze. We focus on run-and-tumble dynamics, and use parameters relevant to *E. coli*, throughout, however the non-dimensional transformations we make should allow the results to be generalised to other types of active particle with motion that can be mapped onto run-and-tumble motion.

5.2 Particle dynamics

5.2.1 Dynamics in free space

The dynamics of the agents in our model are similar to those in Chapter 4. One important difference is that the particles are now points; they do not have a finite volume, and do not interact with one another directly. A second important difference is that the bacteria we considered in the previous chapter did not tumble, as we were modelling a smooth-swimming strain of *E. coli*. In this chapter we attempt to model the behaviour of a typical 'wild-type' strain of *E. coli*, which performs run-and-tumble motion, as outlined in Chapter 2. We do this through an idealised model we outline below.

In our model of run-and-tumble motion, bacteria are represented as self-propelled point particles, which maintain their velocity $\mathbf{v}$, subject to rotational diffusion (during the 'run' phase), until a tumbling event occurs, at which point they instantaneously randomise their orientation and begin another run at the same speed. In reality there is some persistence of direction between runs [71], and there is some evidence that in fact the degree of this persistence itself can have a chemotactic bias [93]. However this is neglected here for the sake of simplicity. Thus, the velocity evolves between timesteps as follows,

$$\mathbf{v}(t+\Delta t) = \begin{cases} R(\eta)\mathbf{v}(t) & & \text{Run} \\ R(\theta)\mathbf{v}(t) & \theta \in [-\pi, \pi] & \text{Tumble} \end{cases}, \tag{5.1}$$

where R is a two-dimensional rotation matrix, η is a random angle picked from a Gaussian distribution in order to model rotational diffusion and θ is a random angle picked from a uniform distribution on the $[-\pi, \pi]$ interval, in order to model the complete reorientation associated with a tumbling event. A particle speed of $\mathrm{v} = 20\,\mu\mathrm{m\,s}^{-1}$, which is realistic for *E. coli* [52], is used.

Given the desire for a minimal model of chemotaxis, the addition of rotational diffusion may be seen as an unnecessary complication, since in free space it is a corrective term in the long-time diffusivity of a swimming, tumbling *E. coli* bacterium, compared to the tumbling behaviour. Experimentally it has been found [133] that $\mathrm{D}_r \simeq 0.1\,\mathrm{rad}^2/\mathrm{s}$, while $\alpha_0 \simeq 1\,\mathrm{s}^{-1}$ [139]. Clearly tumbling is the dominant mechanism of directional relaxation, and it might be expected that

other sources can safely be neglected.

The length-scale of this directional relaxation through tumbling is given by $l_\alpha = v/\alpha_0$. For $v \simeq 20\,\mu\text{m}$, $l_\alpha \simeq 20\,\mu\text{m}$. In the environments considered below, the distance between obstacles is in some cases on this length scale; this means that a particle may not tumble between hitting one wall, and then another. A lack of rotational diffusion, along with the imposition of bacterial alignment parallel with walls means that this would result in narrow 'trains' of bacteria deterministically following particular paths. The high bacterial densities in such regions may cause unrealistically large gradients of chemoattractant. For this reason, we include rotational diffusion in the model, of the experimentally measured magnitude given above.

The third contribution to noise in the bacterial dynamics is explicit translational diffusion, due to Brownian noise. The effective translational diffusion from rotation for the parameters stated is $165\,\mu\text{m}^2/\text{s}$. A typical translational diffusion constant for *E. coli* is $\text{D} = 0.2\,\mu\text{m}^2/\text{s}$. As this is such a small factor in the overall motion of the bacterium, and in the absence of other motivations such as that for explicit rotational diffusion, it is neglected for the sake of simplicity: we set $\text{D} = 0$ throughout.

5.2.2 Chemotaxis

We allow for chemotaxis in the dynamics of these particles by first introducing a measure, the 'chemotactic fitness' of a particle trajectory, $f(\mathbf{v}(t), c(t))$, where $c(t)$ is the chemical concentration at the particle's position at time t.

Secondly, we vary the particle's tumble rate, α, according to f. Unless otherwise indicated in the following, we choose a one-sided response, where the particle's dynamics are identical to those in free space if it is moving in an unfavourable direction (Fig. 5.1). This is the experimentally relevant case for *E. coli* [45, 92, 140]. When the particle is moving in a favourable direction, we choose a linear response:

$$\alpha(f) = \begin{cases} \alpha_0(1-f) & f > 0 \\ \alpha_0 & f \leq 0 \end{cases}, \tag{5.2}$$

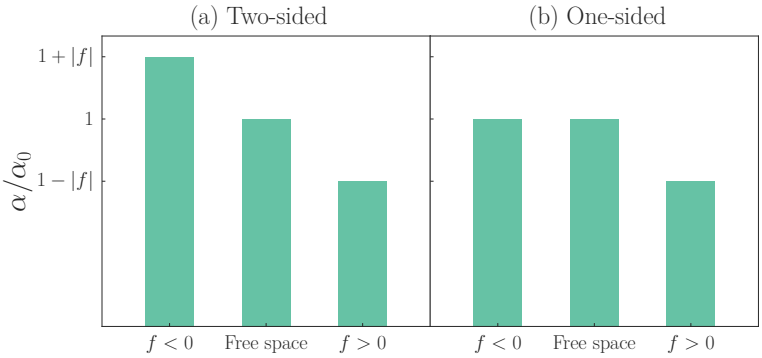


Figure 5.1 *An illustration of how the bacteria's one-sided response is implemented in our simulations. (a) In a two-sided response, if the bacteria is moving in an unfavourable direction, $f<0$, its tumble rate increases; if it is moving in a favourable direction, its tumble rate decreases. (b) In a one-sided response (which mirrors the true behaviour of E. coli), only the favourable direction induces a response in the tumble rate.*

where α_0 is the free-space tumbling rate. In simulations we have chosen $\alpha_0 = 1\,\mathrm{s}^{-1}$, approximately equal to that seen experimentally. This response allows for the possibility of $f<0$, at which point the tumble rate is taken to be zero. Without other sources of noise this may result in agents becoming 'stuck' in particular trajectories, if the chemical gradient is steep enough. However, the presence of rotational diffusion provides a baseline source of rotational noise, such that ergodicity is preserved. Investigation of more physically realistic response functions is a worthwhile direction for future work, however since we are interested in the initial formation of clusters, which is caused by small gradients due to density fluctuations, the form of the function is not expected to qualitatively affect our conclusions. Where indicated, we compare the results using this one-sided linear response to a two-sided linear response, where the tumble rate may both increase and decrease in response to chemical gradients. This shows that the essential conclusions we draw are not substantially affected by the 'sidedness' of the response, beyond shifting the point at which transitions occur by a constant factor.

The form of f we choose is proportional to the instantaneous degree of alignment between the particle's direction and the chemical gradient,

$$\begin{aligned} f &= \chi\frac{\dot{c}}{\mathrm{v}}\,, \\ \dot{c} &= \mathbf{v}(t)\cdot\nabla c(t)\,, \end{aligned} \tag{5.3}$$

$\dot{c}$ represents the time-derivative of the chemical concentration in the particle's

reference frame, and χ controls the strength of the response. In reality, bacteria such as *E. coli* cannot measure $\nabla c(t)$ directly, as the differences in chemical concentration across their cell body are too small to be measured. In fact a complex signalling mechanism is used to reconstruct $\dot{c}$ from absolute local concentration measurements; this is described further in Chapter 6. Although not shown in this chapter, many of the simulations described below were run with particles doing chemotaxis in this more complex fashion; the qualitative results were the same, in terms of confining particle clusters in specific regions, although typical clusters were spatially larger. The idealised measurement method presented here nevertheless maintains the 'perfect adaptation' which is a feature of the more complex model, whereby the absolute concentration of chemoattractant does not affect the bacteria's behaviour; it is only *gradients* in the chemical which are important.

5.3 Chemoattractant dynamics

The concentration of chemoattractant is assumed to have three components to its dynamics: its diffusion through thermal noise; its secretion by the bacteria; and finally its decay, to allow for the existence of a steady-state concentration. This is modelled as a field obeying the following reaction-diffusion equation,

$$\frac{\partial c}{\partial t} = \mathrm{D}_c \nabla^2 c + \phi\rho - \delta c \,, \tag{5.4}$$

where ρ is the bacterial number density and $\mathrm{D_c}$, ϕ and δ are constants quantifying the rate of chemoattractant diffusion, secretion and breakdown, respectively. If c is interpreted in units of area number density of molecules of chemoattractant, and ρ as an area number density of bacteria, then ϕ represents the number of molecules secreted by a single bacterium per unit time. δ represents the fraction of chemoattractant molecules decaying in the same time interval. This model is similar to that used in the classic Keller-Segel model (Eq. (2.1)), with a constant rate of breakdown, and an added secretion term.

For chemoattractant dynamics, the following parameters were used in all simulations: $\mathrm{D_c} = 1000\,\mu\mathrm{m}^2\,\mathrm{s}^{-1}$, $\phi = 1\,\mathrm{s}^{-1}$, $\delta = 0.01\,\mathrm{s}^{-1}$, $c(x,\, t=0) = 0$ and $\rho(x,\, t=0) = \rho_0 = 3.5 \times 10^{-3}\,\mu\mathrm{m}^{-2}$. These correspond to biologically realistic values for a chemoattractant of a molecular weight similar to aspartate, an example

of a molecule to which chemoreceptors of *E. coli* respond [52, 59, 141].

The field ρ is determined from the distribution of particles, by a simple method of counting the number of particles within each unit cell of the lattice. Diffusion of the chemoattractant acts to smooth out numerical fluctuations which might otherwise arise from this method. We use a field cell size of $\Delta x_l = 20\,\mu\text{m}$.

5.4 Results in one dimension

We consider first the simple case where we confine the particles to move in a single dimension, and investigate the interplay between their movement, secretion of the chemoattractant, and their chemotactic response in this environment.

Before we show the results obtained from the agent-based simulations, we present results obtained by treating the bacteria through a drift-diffusion partial differential equation. This allows us to validate our microscopic, agent-based model by comparing it with the coarse model and verifying that both give similar conclusions.

5.4.1 Analytic model

The behaviour of the bacteria may be expected to map onto a drift-diffusion model, where the particles' density diffuses, and has a flux in the direction of increasing c. This is the model used for the bacterial density in the Keller-Segel model, where we choose that the bacterial diffusion and mobility coefficients are independent of the chemoattractant concentration, c. Since our particles are only sensitive to *gradients* in this field, this is the appropriate choice,

$$\frac{\partial \rho}{\partial t} = \mathrm{D}_\rho \nabla^2 \rho - \nabla \cdot (\mu \nabla c \rho) \ , \tag{5.5}$$

where D_ρ is the effective diffusion constant of the particles, and μ their mobility. In this coarse-grained model, these two parameters are the only values which characterise the bacterial dynamics; the detailed method by which bacteria achieve a net drift towards gradients, involving response functions and fitness functions, is not involved here.

In order to simplify analysis of this equation, along with the equation governing

Parameter	Unit	Reduced Parameter	Definition
t	s	$\tilde{t}$	$t\delta$
x	µm	$\tilde{x}$	$x\sqrt{\delta/\mathrm{D}_\rho}$
ρ	$\mu\mathrm{m}^{-2}$	$\tilde{\rho}$	ρ/ρ_0
c	$\mu\mathrm{m}^{-2}$	$\tilde{c}$	$c\delta/(\rho_0\phi)$
D_ρ	$\mu\mathrm{m}^2\,\mathrm{s}^{-1}$	$\tilde{\mathrm{D}}_\rho$	D_ρ/D_c
μ	$\mu\mathrm{m}^4/\mathrm{s}$	$\tilde{\mu}$	$\mu\phi\rho_0/(\mathrm{D}_\rho\delta)$
ρ_0	$\mu\mathrm{m}^{-2}$	$\tilde{\rho}_0$	1

Table 5.1 *Physical parameters used in the simulation, along with their non-dimensional equivalents.*

the dynamics of c (Eq. (5.4)), their variables and parameters were made non-dimensional by the transformations shown in Table 5.1. The reduced units defined here are used in the remainder of the chapter. These transformations result in a simplified version of Eq. (5.4) and Eq. (5.5),

$$\begin{aligned} \frac{\partial\tilde{\rho}}{\partial\tilde{t}} &= \tilde{\mathrm{D}}_\rho\tilde{\nabla}^2\tilde{\rho} - \tilde{\mu}\tilde{\rho}\tilde{\nabla}\cdot\left(\tilde{\rho}\tilde{\nabla}\tilde{c}\right) \\ \frac{\partial\tilde{c}}{\partial\tilde{t}} &= \tilde{\nabla}^2\tilde{c} + \tilde{\rho} - \tilde{c} \end{aligned} . \tag{5.6}$$

It can be seen that a steady-state solution of this equation is $\tilde{\rho} = 1$, $\tilde{c} = 1$. We can investigate the stability of the system about this point by initialising the system in this state, with a small perturbation, and inspecting how the perturbation evolves with time, when the dynamics of c obeys Eq. (5.4). To this end, we initialise the system in the state,

$$\begin{aligned} \tilde{\rho}(\tilde{x}, \tilde{t}=0) &= 1 + \epsilon\sin(2\pi\tilde{x}/\tilde{L}) \\ \tilde{c}(\tilde{x}, \tilde{t}=0) &= 1 + \epsilon\sin(2\pi\tilde{x}/\tilde{L}) \end{aligned} , \tag{5.7}$$

where ϵ is a small quantity characterising the size of the initial perturbation, in this case we used $\epsilon = 1 \times 10^{-5}$. This was run until steady state was reached, for a range of values of $(\tilde{\mathrm{D}}_\rho, \tilde{\mu})$. The stability of the system was determined by computing the variance of $\tilde{\rho}$ across the system at steady state. (The range between maximum and minimum $\tilde{\rho}$ was also used as a measure, with identical conclusions). Figure 5.2 shows a section of parameter space, along with the stability of each point around the above solution. It can be inferred from this that the system becomes unstable when $\tilde{\mu} > \tilde{\mathrm{D}}_\rho$.

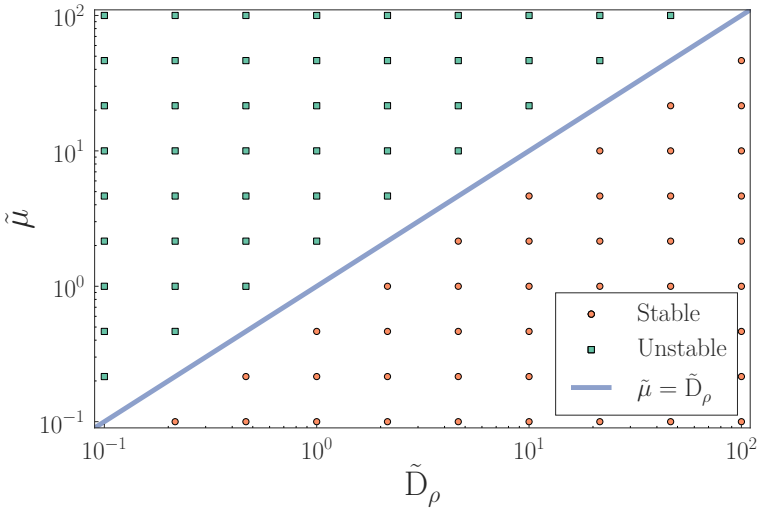


Figure 5.2 *Stability of a uniform distribution of chemotactic bacteria to a small perturbation, when modelled through a drift-diffusion equation, as a function of its two non-dimensional parameters: $\tilde{\mu}$, which represents the mobility of the bacteria relative to their diffusion, and $\tilde{\mathrm{D}}_\rho$, which represents the relative rate of diffusion of the bacteria relative to the chemoattractant molecule. There is a transition to instability at the point where $\tilde{\mu} = \tilde{\mathrm{D}}_\rho$.*

5.4.2 Microscopic model

We now present results also in one dimension, as obtained by representing bacteria through the microscopic model we have described.

For comparison with the coarse-grained analytic model, it is necessary to derive the parameters in Eq. (5.5) from the microscopic parameters. The mobility parameter in one dimension may be derived by considering rates of switching between populations of left- and right-moving populations in a linear gradient at steady state. For a population of n particles, with n_l moving left, and n_r moving right, the change in the right-moving population is,

$$\frac{\partial n_r}{\partial t} = \frac{1}{2}\alpha_l n_l - \frac{1}{2}\alpha_r n_r \,, \tag{5.8}$$

where α_l and α_r are the tumbling rates of left- and right-moving particles, respectively. The factors of a half account for the probability of a tumbling event leaving the particle's direction unchanged. At steady-state, therefore,

$$\alpha_r n_r = \alpha_l n_l \,, \tag{5.9}$$

which, substituting $n_l = n - n_r$, allows us to find the number of right-moving particles,

$$n_r = n\frac{\alpha_l}{\alpha_l + \alpha_r}\,. \tag{5.10}$$

The mobility coefficient represents the drift speed of a particle in the direction of a chemical gradient, per unit gradient. In one dimension,

$$\mu = \bar{\mathrm{v}}/\frac{\partial c}{\partial x}\,, \tag{5.11}$$

where $\bar{\mathrm{v}}$ represents the average speed of a particle in the direction of the gradient. We will assume the gradient increases as we move right, so that for our population above, this average speed is given by,

$$\bar{\mathrm{v}} = \frac{1}{n}\left(n_r \mathrm{v}_0 - n_l \mathrm{v}_0\right)\,, \tag{5.12}$$

which leads, after substituting Eq. (5.10), to

$$\bar{\mathrm{v}} = \mathrm{v}_0\left(2\frac{\alpha_l}{\alpha_l + \alpha_r} - 1\right)\,. \tag{5.13}$$

From Eq. (5.2), $\alpha_r = \alpha_0\,(1 - f_r)$, and $\alpha_l = \alpha_0\,(1 - f_l)$, where f_r and f_l are the chemotactic fitnesses of right- and left-moving particles, respectively. Substituting these values shows that,

$$\bar{\mathrm{v}} = \mathrm{v}_0\left(\frac{2\,(1 - f_l)}{2 - f_l - f_r} - 1\right)\,. \tag{5.14}$$

We will demonstrate the case of the two-sided response, however the one-sided case proceeds similarly. Using our fitness function from Eq. (5.3),

$$\begin{aligned} f_r &= \chi\tfrac{\partial c}{\partial x} \\ f_l &= -\chi\tfrac{\partial c}{\partial x}\,, \end{aligned} \tag{5.15}$$

$$\bar{\mathrm{v}} = \mathrm{v}_0\chi\frac{\partial c}{\partial x}\,, \tag{5.16}$$

which leads finally to,

$$\mu_2 = \mathrm{v}_0\chi\,, \tag{5.17}$$

where the subscript indicates its applicability to the two-sided response. One can similarly show that the one-sided response leads to a mobility parameter $\mu_1 = \mathrm{v}_0\chi/(2-\chi)$.

With regards to the diffusion parameter, we have already derived in Section 3.3.3 that in one dimension, $\mathrm{D}_\rho = \mathrm{v}_0^2/\alpha_0$. In the remainder of the chapter, results for the microscopic model are presented in the non-dimensional equivalents of these two parameters.

It should be noted that implementing slow directional relaxation, like rotational diffusion, in one dimension is not meaningful, as direction is intrinsically discrete – a particle's direction is described by either '-1' or '+1'. As such, rotational diffusion is not implemented in the one-dimensional version of the agent-based model. In two dimensions, problems arise from a lack of rotational noise, since the environment may no longer be ergodically sampled by all particles, *i.e.* their probability of visiting some areas of the system may become zero. For a particle with a constant non-zero speed in one dimension, this is not possible, as all possible trajectories of the particle ('left' or 'right') visit all points in the system, and so the issue is not important.

The distributions of the bacteria within a cluster as predicted by the analytic and microscopic models are in good agreement, Fig. 5.3.

As the response-strength parameter, $\chi \to 0$, particles do not respond at all to chemical gradients, so a uniform distribution of particles is to be expected. As $\chi \to \infty$, particles exhibit a maximal response to chemical gradients, and clustering should be expected. Figure 5.4 shows these two states. It is clear that there is a difference between these situations, but it is useful to introduce an order parameter which quantifies the degree of clustering in the system.

5.4.3 Measuring clustering

There are many algorithms available for clustering a set of points in a space, from the area of data analysis. A disadvantage of several of these is that they require

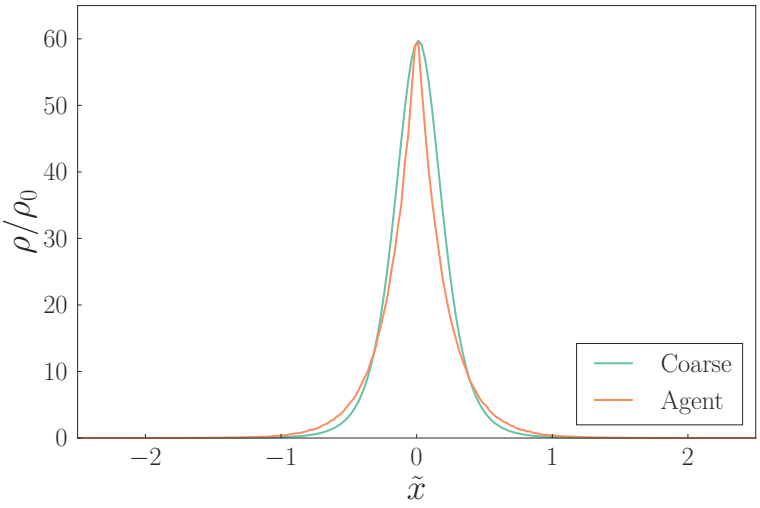


Figure 5.3 *Shape of a bacterial cluster in one dimension, as predicted by agent simulations, and by numerically solving a drift-diffusion equation with parameters inferred from the microscopic parameters of the agent model.*

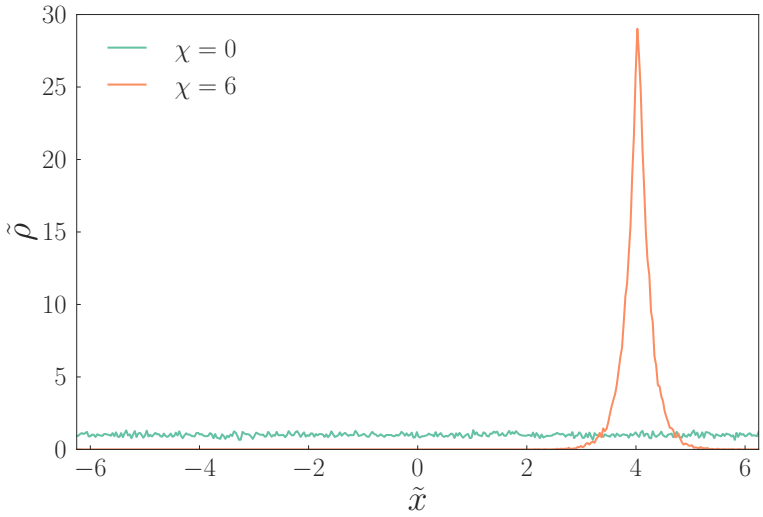


Figure 5.4 *Typical distributions of agents at steady state, binned into a density field, in the uniform (green) and clustered (orange) state.*

prior specification of the number of clusters to assign (such as k-means clustering and spectral clustering), which we do not know; a disadvantage of others is that they do not scale well to thousands of points (affinity propagation, mean shift clustering [142]). One of the few remaining options is agglomerative clustering. In this method, each point is initially considered to be in its own cluster. At each iteration, clusters whose points are 'close enough together' are merged into one.

There are a number of options for how to define 'close enough together', for example, if *all* points from one cluster lie within some distance of *all* points of another, these clusters can be considered to be very close, and merged. This is referred to as the 'complete-linkage' criterion. At the other end, one may require only that *any one* point in one cluster lies within some distance of *any one* point in another — this is referred to as the 'single-linkage' criterion. There are also intermediate cases, such as inspecting the average point separation.

A downside of more rigorous criteria, such as complete-linkage, is that they are slow to calculate, as every pair of points must be compared to confirm that clusters can be merged. The downside of more lax criteria, such as single-linkage, is that they are vulnerable to a small number of 'bridge points' spuriously linking two clusters. This becomes an issue if clusters are not well-separated, which fortunately is not the case for all conditions encountered here. As such the single-linkage criterion was chosen and is used throughout.

It should be noted that since we model the agents in periodic space, the distance is not quite the usual Euclidean distance. In calculating the distance between points, the minimum-image convention is used, where the distance between two points, r_{ij}, is given by,

$$\begin{aligned} x_{ij} &= \left|x_i - x_j\right| \\ x_{ij,p} &= \min(x_{ij}, L_x - x_{ij}) \\ r_{ij,p} &= \sqrt{x_{ij,p}^2 + y_{ij,p}^2 + \dots}. \end{aligned} \tag{5.18}$$

There is a parameter which must be determined before this clustering can be performed: the distance used in the above criterion where clusters are considered 'close enough'. This depends on the typical separation of points in a uniform distribution, and the length scale of clusters when they are present. Too small a value results in no detection of clusters; too large results in false positives. A reasonable value was determined by trial-and-error, by using a sample of uniform

and clustered distributions of agents at the densities used for all simulations, and varying the distance until the intuitively expected results were robustly returned. These values were found to be $r_{c,1\mathrm{D}} = 5\,\mu\mathrm{m}$ in one dimension, and $r_{c,2\mathrm{D}} = 20\,\mu\mathrm{m}$ in two dimensions.

Once the clustering has been performed, the remaining problem is how to characterise the result of the cluster analysis, which is essentially a label assigned to each point indicating to which cluster it belongs, to measure how 'clumpy' the points are, meaning to what extent are their positions positively correlated. To be useful as an order parameter, the measure should return zero in the uniform state, and unity when all particles belong to a single, coherent structure. A first approach might be to simply calculate the fraction of particles that lie in the biggest cluster. This satisfies the conditions just mentioned (assuming the clustering has been done sensibly), however a problem with this measure is that it neglects all information about points not in the largest cluster. In extreme cases this is not important, but when there are several clusters of a meaningful size it can result in counter-intuitive numbers. For example, if there are 1000 particles, distributed in three clusters with populations $\{900, 99, 1\}$, it seems sensible to conclude that this system is in a more strongly clustered state than one with 101 clusters of populations $\{900, 1, 1, 1, 1, \ldots\}$, however the 'biggest cluster fraction' measure would assign these two cases the same value.

An improved, albeit more complicated measure, referred to henceforth as the 'clumpiness' measure, κ, considers the number of points in each cluster, n_i. The measure is defined as,

$$\kappa = \sum_i \frac{n_i}{N} \frac{n_i - 1}{N - 1}\,. \tag{5.19}$$

This measure satisfies the extreme conditions given above, and is robust to corner cases such as those already outlined. For large N, κ is approximately equal to the sum of the squares of the population fractions; the reason for using the more complicated definition can be seen by considering the value of κ when the populations are entirely un-clustered, that is, $\{1, 1, 1, 1, \ldots\}$. In this case, we wish the measure function to return a value of zero, which a simple sum of squares would not yield.

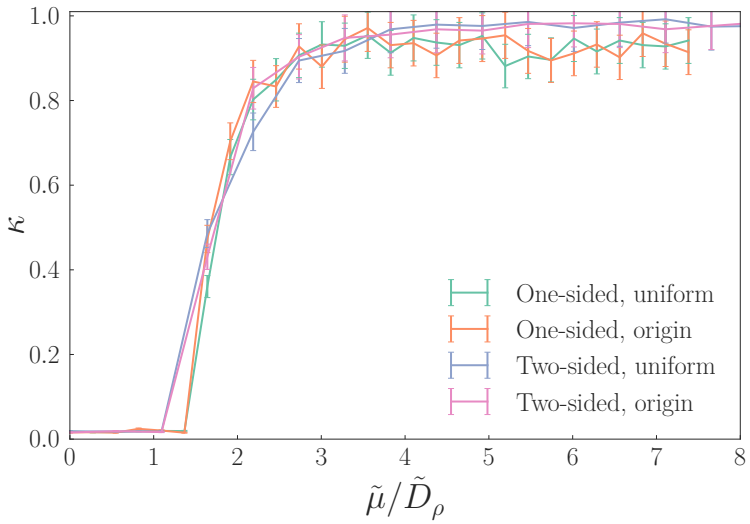


Figure 5.5 *The degree of clustering observed in a one-dimensional system at steady state, for a range of chemotactic sensitivities, χ, expressed as a fraction of the threshold where a uniform distribution is predicted to be unstable in a drift-diffusion model, $\tilde{\mu} = \tilde{\mathrm{D}}_\rho$. The system is run for particles modulating their tumble rate when the gradient is increasing or decreasing (two-sided), or only increasing (one-sided). Data is also shown when initialised from an initially uniform state, and from a delta-function distribution. There is no evidence of hysteresis, as the clustering is independent of whether the system approaches its steady state from an initially clustered or non-clustered state.*

5.4.4 Effect of chemotactic sensitivity and form of response

Figure 5.5 shows the effect of varying the chemotactic sensitivity on the clustering. The results are independent of whether the system is initialised in a clustered or non-clustered state – there is no evidence of history dependence of the system in one dimension. The transition for the two-sided case occurs at $\tilde{\mu} \simeq 1.1\tilde{\mathrm{D}}_\rho$, while for the one-sided at $\tilde{\mu} \simeq 1.5\tilde{\mathrm{D}}_\rho$. The analytic model predicted that instability to cluster formation should occur at $\tilde{\mu} = \tilde{\mathrm{D}}_\rho$. The small deviations from this analytic instability point may be due to numerical errors from the finite time- and space-steps.

5.5 Results in two dimensions

In two dimensions, we find that the analytic drift-diffusion model no longer matches with agent-based simulations. There are a number of reasons for this. The first is that in two dimensions, the minimal model outlined above leads either to no clustering, below the instability transition point, or collapse of the bacterial density into one or more singularities. These points of infinite density are clearly not physical. This can be remedied, and many models have been so derived, by modifying various terms in Eq. (5.5), usually adding a functional form to the mobility coefficient μ. However an issue that would remain is the effect of physical obstacles in the environment. In later sections of the results we consider cases where the environment contains physical obstacles; the time-asymmetric interactions of bacteria with which are difficult to represent through coarse-grained modelling, and so agent-based approaches are required. Figure 5.6 shows the stability of the analytic model in a uniform two-dimensional system, determined through the same numerical method that was carried out in the one-dimensional case, which shows the increased stability of the non-clustered state in two dimensions compared to one dimension, which will be confirmed later on in our agent-based simulations, where a larger chemotactic mobility is required to induce clustering. From now on, however, we focus on results obtained through agent-based simulations.

Figure 5.8 shows the equivalent of Fig. 5.5 but in two dimensions. As in the one-dimensional case, we initialise particles either in a uniform distribution, or in a delta-function distribution, in a uniform space with periodic boundaries, for a particular value of χ, and measure κ at steady-state. The threshold for clustering now depends on the history of the system, *i.e.* if the particles are initialised in a uniform or clustered distribution. For an initially clustered state, the threshold is identical to the one-dimensional case. For an initially uniform state, the threshold is increased by a factor of 3. This is in quantitative disagreement with the 2D stability analysis, which predicted instability of the uniform state at $\tilde{\mu} = 4\tilde{\mathrm{D}}_\rho$; this is likely due to numerical approximations, and the qualitative features are in agreement. It can also be seen that at larger χ, from a uniform distribution, multiple clusters are formed. This difference from only observing single clusters in 1D may be due to the increased space available for clusters: in one dimension, if a cluster is assumed to have a length of 2 in non-dimensional units, the fraction of the system occupied by a cluster is $2/\tilde{L}$. For the 1D simulations, $\tilde{L}_1 \simeq 12$, so a

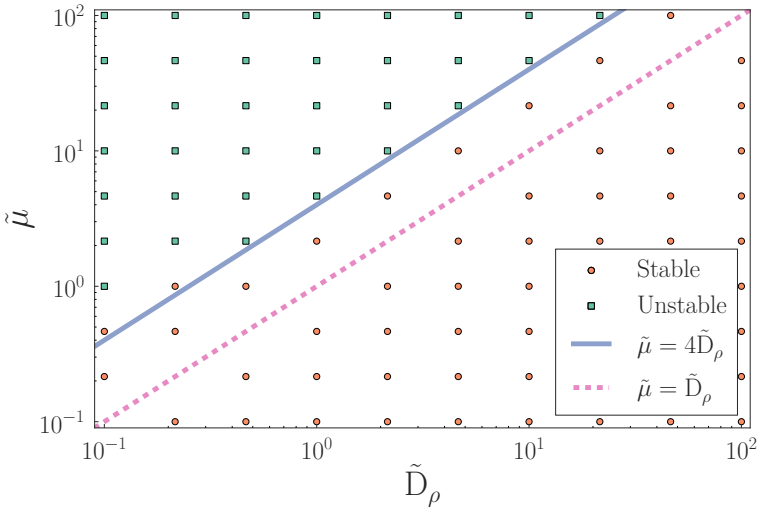


Figure 5.6 *Stability of a uniform distribution of chemotactic bacteria to a small perturbation, when modelled through a drift-diffusion equation, as a function of its parameters $\tilde{\mu}$ and $\tilde{D}_\rho$. There is a transition to instability at the point where $\tilde{\mu} = 4\tilde{D}_\rho$. Also shown is the 1D instability point for comparison. Although increasing the dimensionality stabilises the system, the instability always moves towards a solution with one or more singularities in $\tilde{\rho}$, in contrast to the smooth solutions found in the 1D case.*

cluster occupies around 16 % of the system. In two dimensions, if we approximate a cluster as a circle of diameter 2, a cluster occupies $\pi/\tilde{L}^2$. For the 2D simulations, $\tilde{L}_2 \simeq 18$, meaning a cluster occupies around 1 % of the system. Due to this, there is space for multiple clusters to exist in the system, and at large χ, multiple clusters form and persist at long times. The number of clusters that are formed fluctuates between simulation runs, and since the 'clumpiness' value is smaller for many clusters with small populations than for few clusters with large populations (this can be verified by computing values of κ for different cluster distributions), the measured degree of clustering also fluctuates between simulation runs as different numbers of clusters are randomly formed. Figure 5.7 shows typical configurations in uniform 2D space in both phases. As expected, there is heterogeneity in both ρ and c, with the maximum chemoattractant concentration stable at around 20 times higher than the average concentration. Its length scale is similar to that for the one-dimensional case. The system length in one dimension was chosen to be large enough that gradients generated by a cluster did not self-interact through the periodic boundaries, and was bounded by computational time. In two dimensions, the system length was additionally constrained by the need to accommodate all geometries investigated; in particular, to contain several traps with reasonable spacing between them.

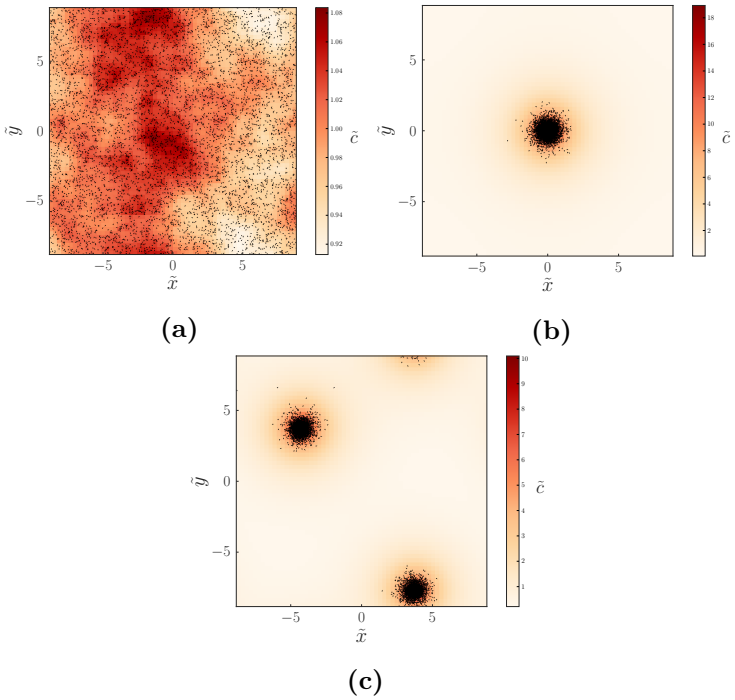


Figure 5.7 *Typical distributions of particles and chemoattractant in a two-dimensional, uniform environment, in various states of clustering. All data is for one-sided chemotaxis. Note that the lengths $\tilde{x}$ and $\tilde{y}$ are in non-dimensional units, where one unit represents the typical length over which the chemoattractant decays. The chemoattractant concentration, $\tilde{c}$ is similarly shown in reduced units, where one unit represents the expected concentration for a uniform distribution of particles at steady-state. Note that the scale of the color-bars are not equal between snapshots, in order to show the fine detail of each. (a) $\chi = 0$, which results in a uniform distribution where $\tilde{c} \simeq 1$ everywhere. (b) $\chi = 1000$, resulting in a single cluster, whose location has been shifted to the centre for the sake of clarity. (c) $\chi = 1600$, At large sensitivities, multiple, stable clusters are formed, of the same size as a single cluster, but necessarily with a reduced density.*

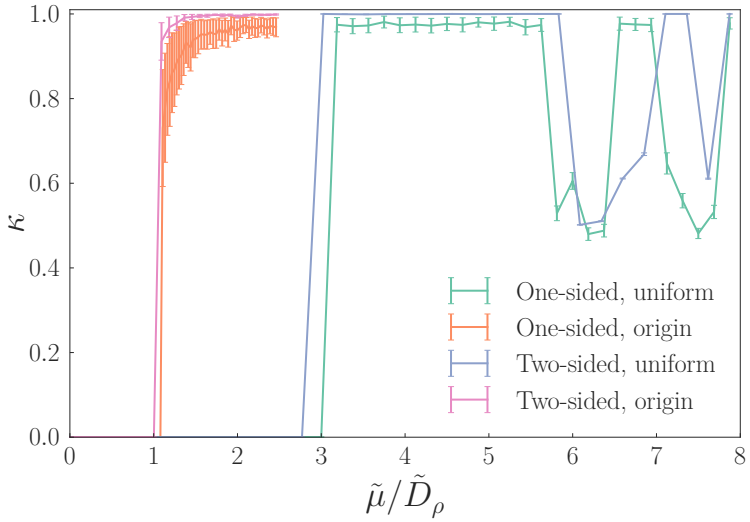


Figure 5.8 *The degree of clustering observed in a two-dimensional system at steady state, for a range of chemotactic sensitivities, χ, expressed as a fraction of the instability point derived previously. The system is run for particles modulating their tumble rate when the gradient is increasing or decreasing (two-sided), or only increasing (one-sided). Data is also shown for initially either uniform bacterial distributions ('uniform'), or those with a delta-function distribution located at the system origin ('origin'). For the initial uniform distributions, the steady-state κ fluctuates at large χ. This is due to the formation of multiple (typically two) stable clusters. The difference in threshold between uniform and delta-function shows clear history dependence, where a cluster, once established, may persist at a value of χ where no clustering would be observed from a uniform distribution.*

5.5.1 Model of a quasi two-dimensional confining environment

Dense clusters of micro-organisms in free space are not unheard of; the algae *Volvox* forms motile colonies containing thousands of members in fresh water [143]. Much more common, however, is the formation of biofilms on surfaces. The advantages of such a living environment are outlined in Chapter 2. For this reason, it is important to consider surfaces when investigating the formation and development of clusters of bacteria. In the following sections we measure the effect of surfaces in controlling the size and location of clusters in an otherwise uniform environment, for particles performing auto-chemotactic behaviour.

Environmental obstacles were incorporated as a two-dimensional lattice of cells, that could either be occupied or not. It should be noted that the dynamics of the particles themselves are nevertheless *off*-lattice throughout the simulations. A bacterium which crosses from a lattice cell that is not occupied, to one that is, indicates the occurrence of a collision with an obstacle. At this point, its position is displaced, and its velocity modified such that it maintains its speed, but is directed parallel to the plane with which it collided (Fig. 5.9),

$$\begin{aligned} \mathrm{v}_{\parallel}(t+\Delta t) &= \mathrm{v}(t) \ , \\ \mathrm{v}_{\perp}(t+\Delta t) &= 0 \ , \end{aligned} \tag{5.20}$$

where $\mathrm{v}_{\parallel}$ and $\mathrm{v}_{\perp}$ are the velocity components parallel and perpendicular to the surface, respectively. Here we also choose for simplicity to preserve the particle's speed. While it has been suggested that only preserving the particle's velocity component parallel to the surface might be a more realistic choice [46, 144], it is not expected that such details will qualitatively affect the results. It is expected that the accumulation near surfaces seen with the 'fully aligning' interaction used here, will be enhanced, with a corresponding enhancement in the effect of obstacles on clustering behaviour.

In this way, three distinct topologies were created and investigated: (a) an environment featuring a single trap-like structure similar in physical dimensions to that used in the experiments of Park *et al.* [66, 68] in order to see how such a structure affects the tendency of the particles to form and sustain clusters; (b) a similar environment featuring multiple traps, in order to introduce multiple

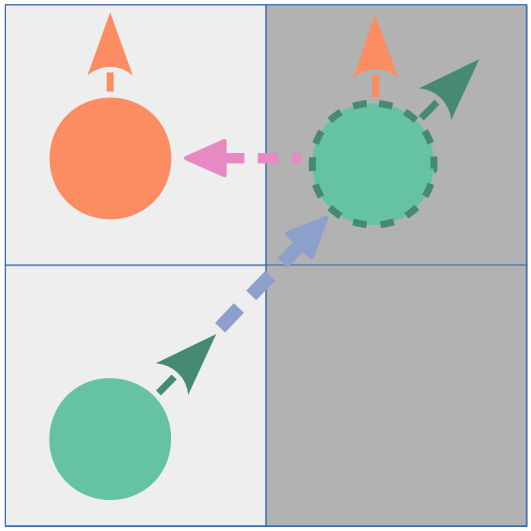


Figure 5.9 *The algorithm used to model velocity alignment of bacteria with surfaces. At each step, a particle's position is displaced without considering the presence of obstacles. If this new position lies within a cell that is marked as obstructed (grey), the velocity component perpendicular to the obstacle's normal vector is set to zero, and the velocity scaled to maintain the previous speed. The position component along the same axis is reset to the previous value. Because the position does not place the particle exactly on the obstacle boundary, this procedure introduces an error in the position of maximum size* $\mathrm{v}\Delta t \simeq 2\,\mu\mathrm{m}$ *every time a collision occurs, however this procedure occurs infrequently for any one particle, and this distance is small compared to all length scales considered.*

'nucleation points' for clustering and (c) a randomly generated maze, again similar to that of Park *et al.* [66, 68] with the aim of investigating cluster formation in highly complex topologies, to see whether and to what extent clusters merge and coarsen at long times.

The parameters used were as follows: all traps have a width, $w = 280\,\mu\text{m}$, with an entrance of width $s = 80\,\mu\text{m}$. The maze channel width was $d = 40\,\mu\text{m}$.

Generation of a random maze

All mazes used in simulations were generated using a modification of the 'Depth-first search' algorithm for traversing graphs (in the computer science sense of the word). This method considers a set of nodes, representing a regular grid of 'rooms', with edges connecting nodes representing the adjacency of the rooms (though not necessarily their connectivity). The algorithm generates the maze by viewing the edges as 'doors', and marking certain doors as being 'open', representing connections between two rooms.

current cell = a random pick from **all cells**
path = {**current cell**}
while path is not empty **do**
    candidate cells = all unvisited cells adjacent to **current cell**
    if candidate cells is not empty **then**
        new cell = a random pick from **candidate cells**
        mark the door connecting **current cell** to **new cell** as 'open'
        current cell = **new cell**
        append **current cell** to **path**
    else
        current cell = the last element of **path**
        remove **current cell** from **path**
    end if
end while

This algorithm forms a maze which is fully connected, meaning that there is a route from any room to any other. This is desirable for our purposes, as we do not want multiple non-interacting clusters in our simulations.

As a minor detail, as stated this algorithm cannot be directly implemented to generate obstacles of the sort used in our model. The algorithm considers cell *faces*

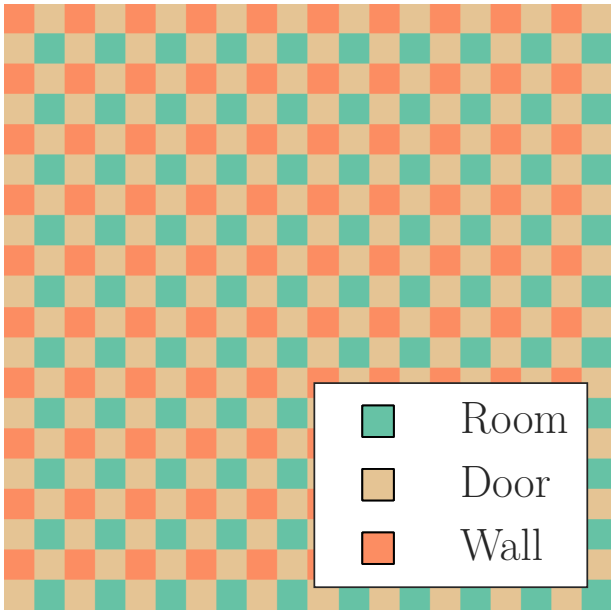


Figure 5.10 *Labelling method of a lattice to allow the implementation of the Depth-first search algorithm to generate a random maze. The algorithm begins with an entirely filled lattice. It then hops between 'room' cells, marking them as empty, as well as the intervening 'door' cell. 'Wall' cells are never visited.*

which are either 'open' or 'closed', while our model considers square cells which are either 'filled' or 'empty' (there is no concept of cell faces). This problem may be overcome by viewing the lattice of cells as overlaid with a pattern as shown in Fig. 5.10, so as to imitate the room-door approach. When the algorithm requires us to mark a 'door' as 'open', and to move to the adjacent room, in our case we mark two cells as 'empty': the 'door' cell which represents the door between adjacent rooms, and also the 'room' cell to which we are moving. Because we always move in steps of two cells, between rooms, the 'wall' cells are never visited, and are always 'filled' when the algorithm terminates.

5.5.2 Results in confining environments

When a trap is placed in the environment, even with no chemotactic response ($\chi = 0$), there is an excess density in the trap's interior. This is due to the time-asymmetric way in which the particles interact with the wall. Figure 5.11 shows the size of this effect by measuring the excess accumulation inside the trap, above what would be expected for a uniform distribution of particles. This is

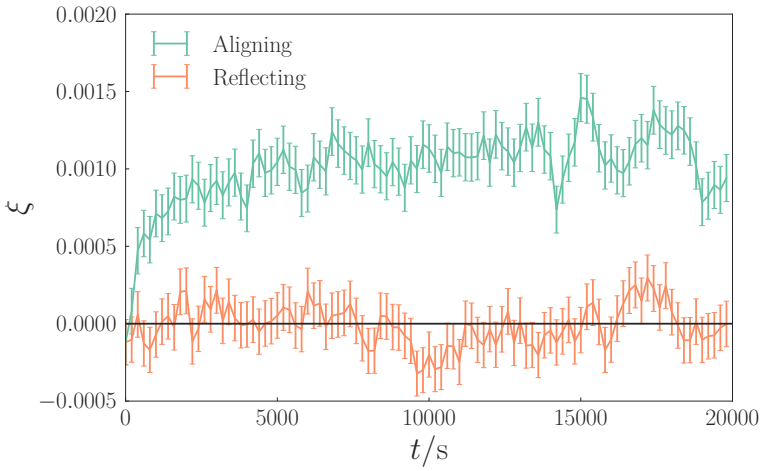


Figure 5.11 *The 'confinedness', ξ, inside a single trap over time of non-chemotactic particles, for two different types of interaction with planar surfaces. For reflecting interactions (meaning that the velocity component perpendicular to the surface normal has its sign inverted) (orange), there is no excess of particles inside the trap, while for aligning interactions (meaning that the perpendicular velocity is set to zero), there is a significant excess. From the dimensions of the trap it can be computed that $\Psi_0 = 1.3\,\%$, which allows the excess density inside the trap to be calculated for the aligning interaction, using Eq.* (5.21)*. For $\xi = 0.0011$, $\Psi \simeq 1.08\Psi_0$, meaning there is an excess of around $8\,\%$.*

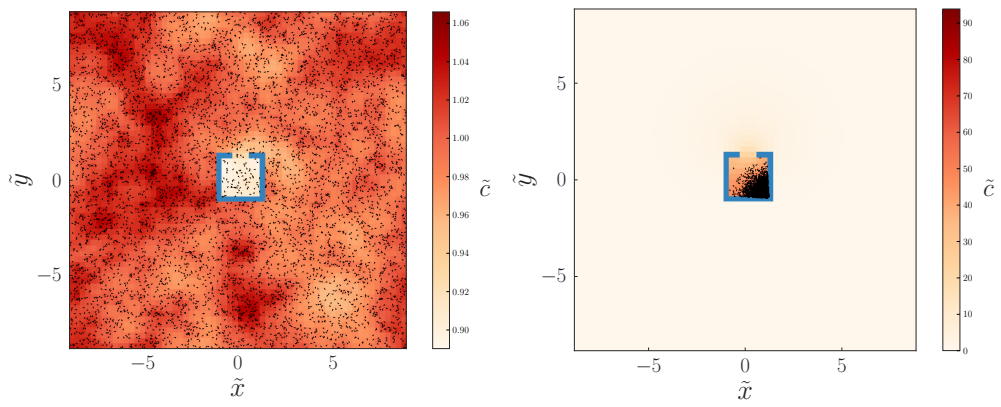
measured through what we call the 'confinedness',

$$\xi = \frac{\Psi - \Psi_0}{1.0 - \Psi_0} \tag{5.21}$$

where Ψ is the fraction of particles inside the trap, and Ψ_0 is the fraction expected for a uniform distribution of particles. This definition is chosen so that $\xi \to 1$ as the particles are entirely confined in the trap ($\Psi \to 1$).

it is important to note that there would be an increased density near an obstacle that was simply a line segment, due solely to the asymmetric wall interaction. This effect can be observed on the trap's *exterior*, where clusters also preferentially form for large χ: see Fig. 5.15c (although with much less preference than for the trap's interior).

The increased density in the trap is larger than that seen near a line segment, however, which is due to the relative length scales of the trap's entrance width, s, and the persistence length of an agent, l_r. An agent moving in free space will randomise its direction on average after travelling l_r. If an agent is moving parallel



(a) **(b)**

Figure 5.12 *Typical distributions of particles and chemoattractant in a two-dimensional environment containing a single trap, without chemotaxis, and with sufficiently high χ to induce clustering. (a) $\chi = 0$. Without chemotaxis, the trap has no significant effect on the particle distribution. In this snapshot the chemoattractant concentration is around 10 % below its expected value for a uniform distribution, however this is due to density fluctuations due to the finite number of agents. (b) $\chi = 686$. The particles consistently preferentially cluster in the trap interior; specifically, in a single corner opposite the entrance (there is no preference of corner, as expected from the trap's symmetry). Inhibited diffusion of the chemoattractant out of the trap leads to peak concentrations around 5 times higher than in a uniform environment.*

to the trap entrance, it effectively cannot 'see' the entrance for this distance, and so is less likely to leave the trap. For our parameters, $s = 120\,\mu\text{m}$, and $l_r = 20\,\mu\text{m}$, so this effect of the persistence length is large enough to be significant. It is this effect that leads to a preferential upwards density fluctuation in the trap interior, and so preferential cluster formation.

The effect of the trap for particles doing chemotaxis can be seen in Fig. 5.12. From an initial uniform distribution, for $\tilde{\mu}/\tilde{\mathrm{D}}_\rho \lesssim 4$, an approximately uniform distribution is obtained (aside from fluctuations on the order of 10 % due to the aforementioned effects). At sufficiently large χ, $4 \lesssim \tilde{\mu}/\tilde{\mathrm{D}}_\rho \lesssim 8$ a single cluster is formed in a randomly chosen corner in the trap opposite the entrance. For $\tilde{\mu}/\tilde{\mathrm{D}}_\rho \gtrsim 8$, it becomes possible that additional clusters are formed outside of the trap. Once a single cluster has been formed inside the trap, this is stable even when the chemotactic sensitivity is reduced such that $\tilde{\mu}/\tilde{\mathrm{D}}_\rho \simeq 0.3$.

The importance of the trap slit width for the clustering is shown in Fig. 5.14. Although it may seem that narrowing the entrance would make it harder for

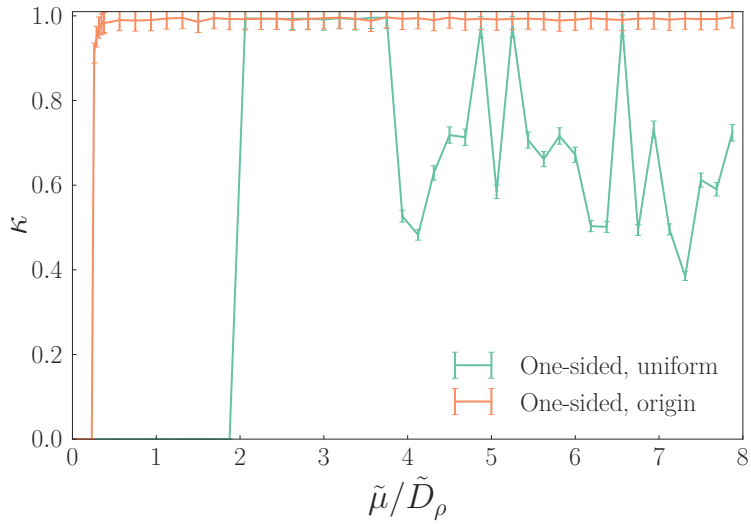


Figure 5.13 *The degree of clustering in a two-dimensional system containing a single trap-like structure, as a function of the chemotactic sensitivity. Considering both sets of data, from the initially uniform or delta-function state, compared with a uniform two-dimensional environment, the threshold for clustering is reduced to around two thirds of the uniform value.*

particles to accumulate in the trap, in fact the opposite is found: a narrower trap allows accumulation to occur at a lower χ.

Although it is clear that a trap can act as a nucleation point for clustering, it is unclear how the particles will respond when multiple equivalent nucleation points are available. To this end, simulations were run in a system featuring five traps of equal dimensions (Fig. 5.15). As can be seen from Fig. 5.18, it is possible for a single cluster to be formed ($\kappa = 1$), however this is less likely than in the single-trap case. In addition, the behaviour is more predictable at higher χ, where more of the traps are occupied as χ is increased. In situations where at an equivalent χ clusters would form at random locations outside the trap, these are now predictably found inside a subset of the traps.

The final environment in which the particles' behaviour is probed is in the maze generated through the algorithm described in Section 5.5.1. We choose a maze wall width of 40 µm. With no chemotaxis, $\chi = 0$, we observe significant density variations due to wall interactions, such that densities higher by a factor of 2.5, and lower by a factor of 2 are generated (Fig. 5.16a). The degree of variation is larger than in the trap environments as the mean distance a particle can travel before colliding with a wall is shorter in the maze, which increases the importance of wall interactions relative to tumbling. The locations of increased density are

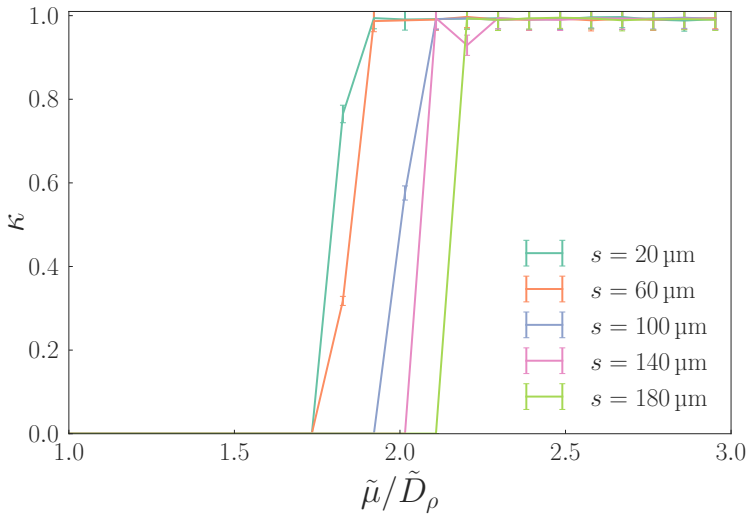


Figure 5.14 *The degree of clustering in a two-dimensional system containing a single trap-like structure, as a function of the chemotactic sensitivity, comparing behaviour with a range of trap entrance widths, s. A smaller s reduces the threshold of chemotactic sensitivity required for clustering inside the trap. The agents here were initialised in a uniform state.*

typically 'dead-ends', which are similar to the trap-like structure in the previously studied environments.

For $\chi > 0$, a large number of clusters form, typically on the order of 10 (Fig. 5.16b). These are stable at long times, and preferentially form in the dead-end structures where increased density is observed without chemotaxis, as expected. A particular feature of the locations where clusters form is the existence of closed 'loops', which are regions where the aligning interaction causes bacteria to repeat a short path for a long time in the absence of tumbling and rotational diffusion Fig. 5.17.

By inspecting Fig. 5.18, we can see that for $\tilde{\mu}/\tilde{D}_\rho \lesssim 2$, the maze enhances the degree of clustering observed, when compared to the most cluster-promoting environment seen so far (that with multiple traps). However, for $\tilde{\mu}/\tilde{D}_\rho \gtrsim 2$, the tendency of the particles to form multiple clusters reduces their ability to coalesce into a single large cluster, which reduces the value of κ. As χ increases, the clusters increase in number, and reduce in density, in a continuous fashion.

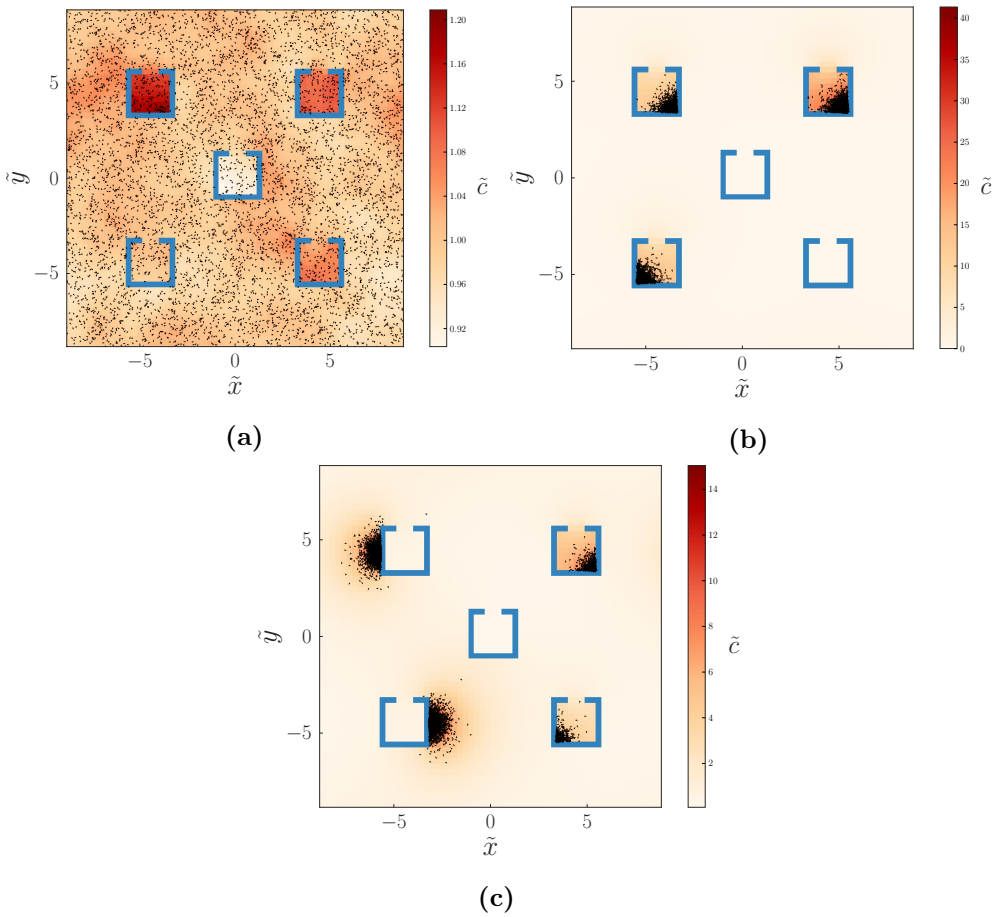
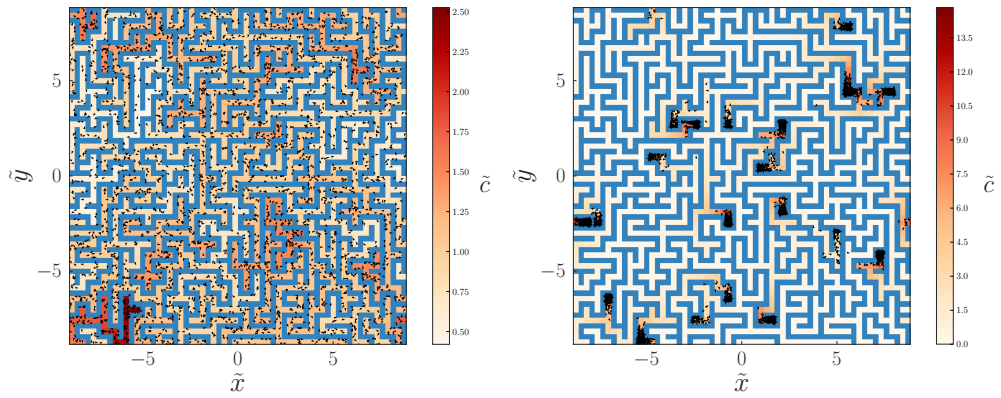


Figure 5.15 *Typical distributions of particles and chemoattractant in a two-dimensional environment containing five traps, in various states of clustering. (a) $\chi = 0$. It can be seen that both upwards and downwards fluctuations in chemoattractant concentration are seen inside the traps, as inhibited diffusion leads to longer-lasting fluctuations. (b) $\chi = 686$. Clusters form in a number of traps at long times. (c) $\chi = 1600$. At large χ, density fluctuations caused by the surface outside of the trap becomes sufficient to induce clustering.*



(a) **(b)**

Figure 5.16 *Typical distributions of autochemotactic particles in a randomly generated maze, in the non-clustered and clustered states. (a) $\chi = 0$. Density variations of up to a factor of 2.5 are seen, caused by wall interactions. (b) $\chi = 800$. The many clusters that are shown are stable at long times.*

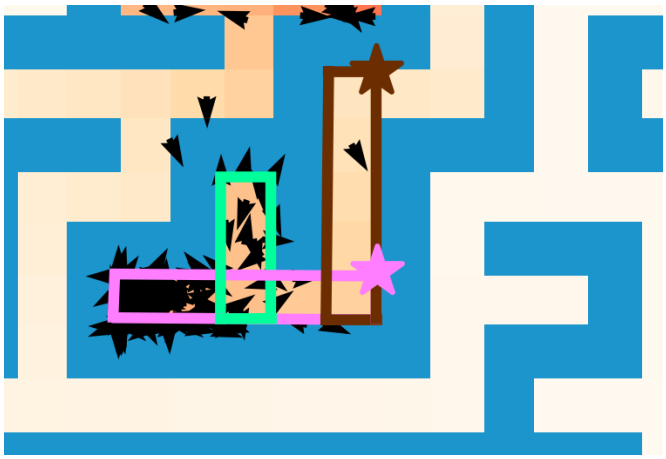


Figure 5.17 *An example of a hierarchy of loops that trap self-propelled particles which align with surfaces. A particle which is moving on the pink loop in a clock-wise direction, has a probability of one half to repeat the loop again when it reaches the point indicated by the star. If it leaves the pink loop, the particle is in the same situation in the brown loop, which may lead back into the pink loop. This leads to a trapping effect which promotes the forming of clusters within such loops. The green loop is an example of a loop which does not exhibit this trapping effect.*

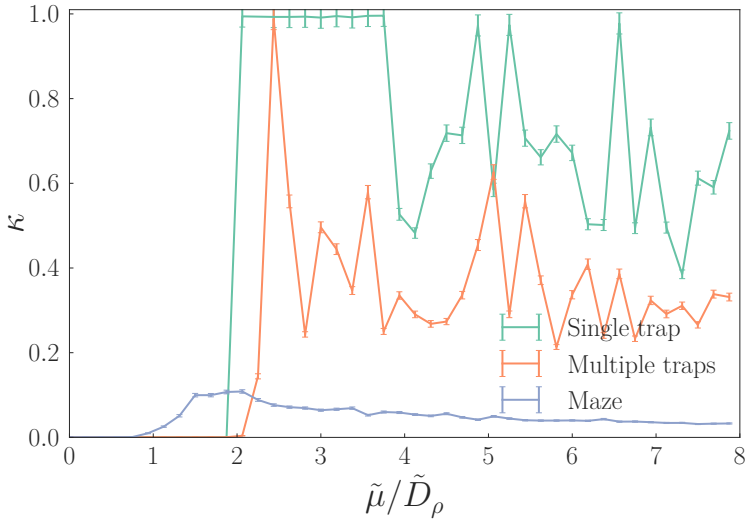


Figure 5.18 *The degree of clustering in a two-dimensional system containing multiple nucleation points for clustering, as a function of the chemotactic sensitivity. With five traps, the threshold can be seen to be similar to in the single-trap case, but the typical degree of clustering is reduced, as there is a tendency for multiple clusters to form in more than one trap. With a maze, the typical degree of clustering is also much reduced, again due to many clusters being formed. In addition, the threshold at which a uniform state is lost is reduced. This is likely because the maze introduces a range of length scales into the system, so that the optimal length scale of confinement is likely to be available in the system, regardless of the particles' dynamics.*

5.6 Conclusion

These results show how microfabricated environments can be used to control the distribution of bacteria that spontaneously aggregate into clusters. One method by which clustering may be controlled is that, rather than attempting to suppress their formation, instead it may be possible to prompt them to arise in predictable locations, chosen to minimise their harm, or maximise their benefit, depending on the biological context. An example of this is the trap-like structures we have described. The optimal shape of such structures has been investigated, showing that those which are difficult for bacteria to enter may counter-intuitively do a better job of attracting clusters.

A second method is to pattern surfaces to prevent the large density fluctuations that have been shown to lead to cluster formation, as demonstrated by the ability to predict the location of chemotactic clusters in a maze-like environment, from the regions of increased density when no chemotaxis is present.

A third, intriguing approach is to design environments which present a large number of potential nucleation points at which particles may form clusters, resulting in many clusters whose population is small. In contexts where clusters are unwanted, such as when the suppression of biofilms is desired, a large number of small clusters may be less harmful than a small number of high-density clusters, if the threshold at which quorum sensing is triggered is not met.

Chapter 6

Migration of active particles in porous media

6.1 Introduction

In the previous chapter we investigated the behaviour of bacteria in artificial environments that are made in the lab. These environments consist of planar surfaces, arranged in various configurations. In many real-world environments, such as in soil, the environment is much more heterogeneous than this unrealistic model world. In this chapter we attempt to model the behaviour of bacteria in more realistic environments, with particular focus on the interaction between the ability of bacteria to estimate spatial gradients, and their directional persistence time.

In free space, this persistence time is bounded from above by rotational diffusion from Brownian motion. However, in complex media, interactions with surfaces must also be taken into account. If the 'mean-free-path' of a ballistic trajectory becomes comparable to the persistence length expected from rotational diffusion, the diffusive characteristics of the bacteria are expected to be dictated less by Brownian dynamics, and more by the interaction of the bacteria with the surfaces when a collision occurs.

This is a biologically realistic scenario, where it is common to find environments with porosities of 20 % to 40 % [145] and with pore sizes from 10 µm to 100 µm [146]. Chemotactic bacteria inhabit such environments. For example, many strains

migrate towards plant roots leading to their promoted growth and disease protection, such as *Pseudomonas fluorescens* towards tomato plants [147], and *Azospirillum brasilense* towards major global agricultural crops [148]. The effect of the structure of a porous environment on these processes is an important factor in understanding and controlling the phenomena.

In the work below, we again focus on chemotaxis through modulated rotational noise. Within this class of methods of achieving a net drift in favourable directions, there are several choices that can be made. Understanding and quantifying the impact of each of these is the aim of this work. As mentioned previously, it is experimentally found that bacteria such as *E. coli* respond to chemotactic gradients in a one-sided fashion, decreasing their rotational diffusion constant when travelling in favourable directions, but not doing the inverse in unfavourable circumstances. This modulation is driven by a mechanism of internal signalling that has feedbacks allowing for adaptation to changing ambient conditions, and averages chemical concentrations over the recent past to estimate how optimal is the current path. The alignment of flagellated bacteria parallel with obstacles is also likely to be an important feature in porous media.

Four core features of the design of prototypical prokaryotic chemotactic dynamics are thus investigated below:

- Rotational noise: discrete (tumbling) or continuous (rotational diffusion)
- Chemotactic response: one- or two-sided
- Chemotactic measurement method: direct spatial measurement or temporal averaging
- Interaction with surfaces: aligning, reflecting, reversing or stalling.

6.2 Particle dynamics

In free space, there are two models we investigate. The first are what are commonly referred to as 'active Brownian particles'. These are self-propelled particles, which move at a constant speed, and undergo rotational diffusion of their direction. The second are run-and-tumble particles. The difference between them is in how their direction is randomised: in the former case, the direction slowly relaxes, while the latter does the same through a discrete switching process.

Whichever method of implementing rotation is employed, both can be modulated to achieve a drift in the presence of chemical gradients. In Chapter 5 we outlined a linear response for tumbling particles,

$$\alpha(f) = \alpha_0(1 - f) \tag{6.1}$$

where f is the 'chemotactic fitness'. We investigate below the behaviour of active Brownian particles with the same linear response, but whose parameter to be modulated is D_r, rather than α.

The second factor of this response which we extend from Chapter 5 is the form of f. There, f is defined as the alignment of the particle with the chemical gradient,

$$\begin{aligned} f &= \chi \frac{\dot{c}}{\mathrm{v}}\,, \\ \dot{c} &= \mathbf{v}(t) \cdot \nabla c(t)\,. \end{aligned} \tag{6.2}$$

In practice, prokaryotes such as bacteria are almost always too small to sense such gradients directly (with potential, rare exceptions [149]). Instead, they proceed through a method of temporal averaging to estimate this quantity. The details of the implementation of this mechanism are outlined below.

6.2.1 Temporal gradient sensing

Biological mechanism

Before introducing the model of temporal sensing that is implemented in our simulations, we outline the biological mechanism which we attempt to mimick. We give the example for *E. coli* as it is this organism whose physical parameters we have used, but other species use similar internal mechanisms of signalling and response (though the details often vary).

Distributed on the surface of the bacterium are specialised proteins which straddle its surface and act as receptors [150]. These are known as 'methyl-accepting chemotaxis proteins' (MCPs) [151]. The external part of these MCPs bind with a particular ligand, which is the molecule to which the chemotactic response is directed. The internal part also binds with a particular signalling protein, known

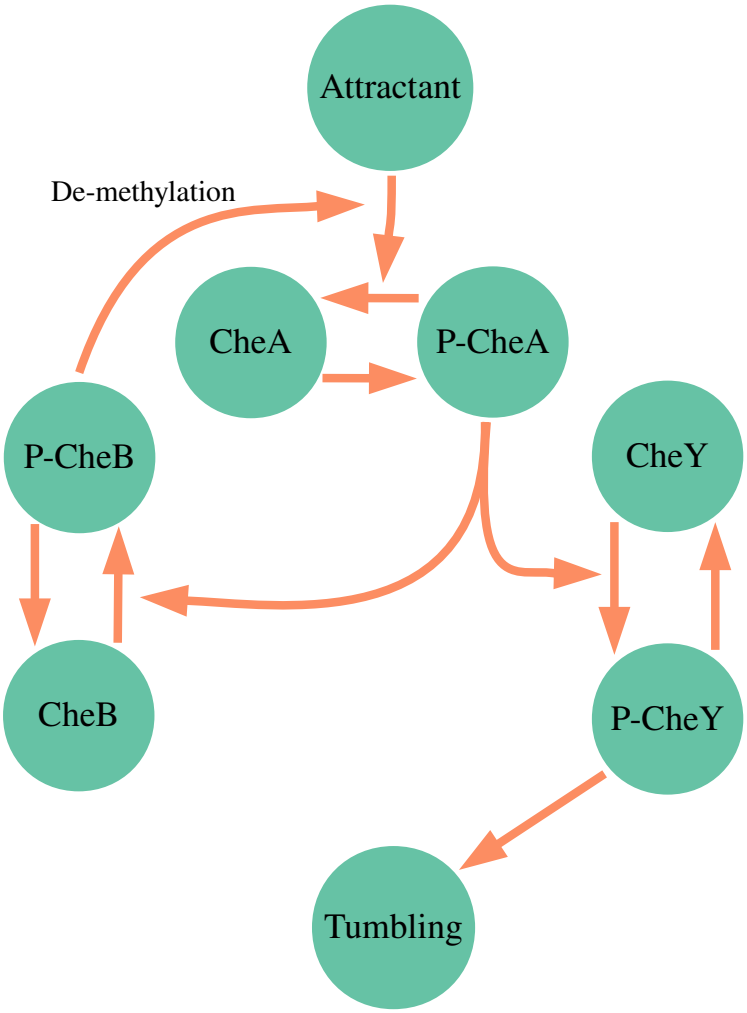


Figure 6.1 *A simplified signalling network representing an adaptive sensing mechanism for the spatial gradient of a chemical attractant. The right-hand path represents the reduced tumble rate in response to increasing attractant concentration; the left-hand path represents the long-term adaptation which allows the bacterium to maintain a consistent base tumble rate, and sensitivity, in environments with differing average attractant concentrations.*

as ‘CheA’ [152], and transfers a phosphate group to it, (‘P-CheA’: phosphorylated proteins will hereafter be referred to with this nomenclature). This phosphorylation is a common way of marking a signalling protein as ‘active’ [153].

In the case of chemical attractants, binding of the receptor with the molecule leads to an inhibition of the rate of activation of CheA. This is the example we will use, however the same logic applies to repellents, except that binding with repellents *promotes* activation of CheA [154].

There are two pathways relating to these MCPs that are relevant to the chemotactic response. The first relates to the response itself; the second to its adaptation to changing ambient conditions.

The response itself is mediated by a signalling protein, ‘CheY’. This is activated by P-CheA, to form P-CheY. This activated form of the protein interacts with the flagellar machinery, acting to reverse its direction from the usual counter-clockwise rotation, which leads to smooth-swimming (a run), to a clockwise rotation, which leads to unbundling of the flagella, and random rotation (a tumble). Therefore an increase in the level of attractant reduces P-CheA, which reduces P-CheY, which reduces the rate of tumbling.

The adaptation is mediated by a signalling protein, ‘CheB’. This is also activated by P-CheA, to form P-CheB. This activated form of the protein removes methyl groups from the receptor. When an MCP is fully methylated, it stops responding to the attractant. This leads to adaptation as follows: if the level of an attractant increases, initially the level of P-CheA (and, therefore, P-CheY and P-CheB) will remain low and the cell will swim smoothly. If the level of attractant remains high, eventually, because P-CheB is not present to de-methylate the receptor, the MCP will stop responding, so the level of P-CheA will increase again. P-CheY will increase, so the cell will tumble again. P-CheB will increase, so the MCPs will be demethylated, and the receptors can once again respond to attractants.

These feedbacks are summarised in Fig. 6.1. The essential outcome is that the bacterium compares the recent concentration with the average over the recent past, and if it is higher than average then the tumble rate is decreased.

The history-dependence produced by the feedback mechanism above implies that the important object controlling the bacteria's tumble rate is a function representing the bacterium's measurement of the local chemical concentration, $c(t)$. So we require a map from $c(t)$ to α. This can be represented by the following functional,

$$\tilde{\dot{c}}(t) = \frac{\mathrm{N}}{\tau^2}\int_0^t K(t-t')c(t')\,\mathrm{d}t'\,, \tag{6.3}$$

where $\tilde{\dot{c}}(t)$ is the bacteria's estimate of its local temporal concentration change at a time t, τ represents the time over which the concentration is averaged, $c(t')$ is the bacterium's measured local concentration at a time t', and $K(t)$ is the kernel which converts the bacterial chemical memory into the estimate.

We can investigate the condition for Eq. (6.3) to be independent of the ambient concentration over long times by decomposing the memory into ambient and corrective terms,

$$c(t) = c_0 + \tilde{c}(t)\,, \tag{6.4}$$

Substituting Eq. (6.4) into Eq. (6.3), and taking $t \to \infty$,

$$\begin{aligned}\tilde{\dot{c}} &= \frac{\mathrm{N}}{\tau^2}\int_0^\infty K(t-t')\left(c_0 + \tilde{c}(t)\right)\mathrm{d}t' \\ &= \frac{\mathrm{N}}{\tau^2}\left(c_0\int_0^\infty K(t-t')\,\mathrm{d}t' + \int_0^\infty K(t-t')\tilde{c}(t)\,\mathrm{d}t'\right)\,.\end{aligned} \tag{6.5}$$

For a perfectly adaptive response,

$$\int_0^\infty K(t)\,\mathrm{d}t = 0\,. \tag{6.6}$$

The kernel we use (which has this property) is one derived by Clark and Grant [155],

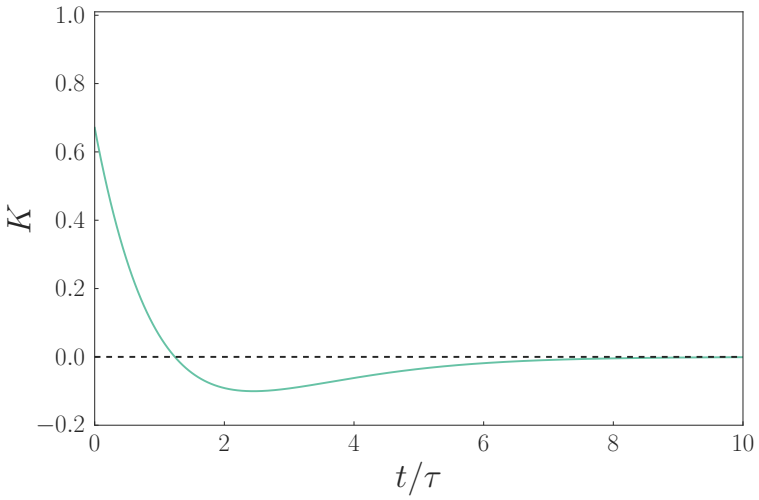


Figure 6.2 *The form of the kernel in Eq.* (6.7)*, derived by Clark and Grant [155]. τ is a time representing the time over which the gradient is averaged.*

and has the form,

$$K(t) = \exp\left(-\frac{t}{\tau}\right)\left(1 - \frac{1}{2}\left(\frac{t}{\tau} + \frac{1}{2}\left(\frac{t}{\tau}\right)^2\right)\right), \tag{6.7}$$

(see Fig. 6.2). To implement temporal chemotaxis in our framework, we simply replace $\dot{c}$ in Eq. (6.2) with $\tilde{\dot{c}}$ as defined in Eqs. (6.3) and (6.7).

6.2.2 Calibration of the two models

Both spatial and temporal models are based on an estimate of the temporal concentration change, and as such have the same parameters. However it should not be expected that the models will necessarily give the same results at equal χ. In order to make the models meaningfully comparable, we calibrate the models so that in at least one regime, the models give the same results. The regime we choose is that where the particle swims in a constant chemical gradient in a constant direction. We find the normalisation constant N in Eq. (6.3) such that in this regime, $\dot{c} = \tilde{\dot{c}}$.

Assume a constant spatial gradient $\nabla c(\mathbf{x}, t) = \nabla c$. For spatial chemotaxis,

$$\dot{c} = \mathbf{v} \cdot \nabla c \tag{6.8}$$

Since the particle is at a constant velocity in a constant gradient, $c(t) = c(0)+t\mathbf{v}\cdot\nabla c$. As shown in the previous section, constant terms vanish for our choice of kernel, so we can ignore the initial concentration term. Substituting this into Eq. (6.3),

$$\tilde{\dot{c}} = \frac{\mathrm{N}}{\tau^2}\mathbf{v}\cdot\nabla c\int_0^t K(t-t')t'\,\mathrm{d}t'\,. \tag{6.9}$$

We can now evaluate the integral (labelled I for convenience) using Eq. (6.7),

$$I = \int_0^t \exp\left(-\frac{t-t'}{\tau}\right)\left(1-\frac{1}{2}\left(\frac{t-t'}{\tau}+\frac{1}{2}\left(\frac{t-t'}{\tau}\right)^2\right)\right)t'\,\mathrm{d}t' \tag{6.10}$$

defining $\hat{t} = \dfrac{t'-t}{\tau}$,

$$I = \tau^2\int_{-t}^0 \exp\left(\hat{t}\right)\left(1-\frac{1}{2}\left(-\hat{t}+\frac{1}{2}\hat{t}^2\right)\right)\left(\hat{t}+t\right)\mathrm{d}\hat{t} \tag{6.11}$$

t is a constant, so this term vanishes, leaving,

$$I = \tau^2\int_{-t}^0 \exp\left(\hat{t}\right)\left(1-\frac{1}{2}\left(-\hat{t}+\frac{1}{2}\hat{t}^2\right)\right)\hat{t}\,\mathrm{d}\hat{t} \tag{6.12}$$

This can be directly computed, giving the result $I = \tau^2\dfrac{3}{2}$. Therefore,

$$\begin{aligned}\tilde{\dot{c}} &= \frac{3}{2}\mathrm{N}\mathbf{v}\cdot\nabla c\\ &= \frac{3}{2}\mathrm{N}\dot{c}\,,\end{aligned} \tag{6.13}$$

so for the models to be comparable, $\mathrm{N} = \dfrac{2}{3}$.

6.2.3 Parameters

The physical parameters we choose for the particles are $v = 20\,\mu\mathrm{m\,s}^{-1}$, and the noise parameter, $D_{r,0} = \alpha_0 = 1\,\mathrm{s}^{-1}$. As the particles do not interact with each

other, the number of particles is only a simulation parameter; we choose n = 5000, and run all simulations up to $t = 1000\,\mathrm{s}$, which was tested as being sufficient to reach steady-state for the highest packing fractions used. A time-step of $\Delta t = 0.01\,\mathrm{s}$ was used. The chemical memory was truncated at an upper bound of $t_{\mathrm{mem}} = 10\,\mathrm{s}$, which inspection of Fig. 6.2 shows is a good approximation. The negative portion of the truncated kernel was re-scaled by a constant factor to preserve the overall zero-area required for perfect adaptation, which required a rescaling factor of around 1.1. This truncated time implies that the chemical memory for a single bacterium should contain $t_{\mathrm{mem}}/\Delta t = 1000$ values. Due to the large memory requirements of storing each bacterium's chemical memory in this way, an implementaton optimization was made, whereby the chemical memory was only updated every $\Delta t_{\mathrm{mem}} = 10\Delta t = 0.1\,\mathrm{s}$; tests comparing the dependence of the free-space drift speed on Δt_{mem} showed that this optimisation was valid.

6.3 Results in a uniform environment

We begin by simply verifying that at times longer than a few multiples of the directional persistence time, tumbling and rotationally diffusing particles diffuse at equal rates in free space, Fig. 6.3. The quantity we are measuring is the slope of the square of the displacement of a particle's position at long times, as a result of its self propulsion with rotational noise (Fig. 6.4).

We then place the particles in a linear chemical gradient, $c(\mathbf{r}, t) = r_x$, and measure their drift speed as a function of their chemotactic sensitivity, χ. The drift speeds in one dimension of particles doing spatial chemotaxis matches the predictions mentioned in Section 5.4.1.

In order to derive analytically the drift speed of these particles in two dimensions, the steady-state probability distribution of orientations must be known. From this perspective, it is useful to view rotational diffusion as the translational diffusion of an overdamped particle in a heterogeneous medium, with periodic boundaries. In this case, the particle's position in fact represents a particular agent's orientation angle, and the periodic boundaries lie at $\pm\pi$. The derivation of such a probability distribution is shown in Appendix A, with the following prediction for the drift

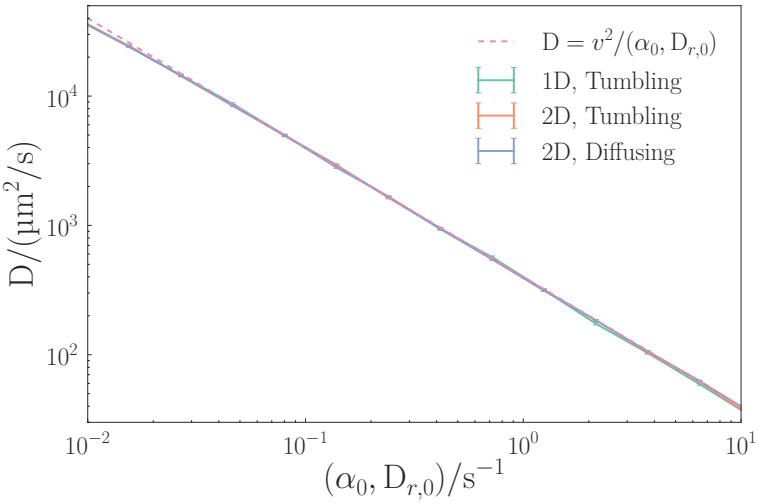


Figure 6.3 *Translational diffusion over long times as a function of the noise parameter, for particles whose direction relaxes through rotational diffusion (*ABP*s) or through tumbling (*RTP*s), in one and two dimensions. Rotational diffusion is not possible in one dimension as direction is a discrete quantity. In all cases, there is one degree of freedom, so the diffusion matches the theoretical prediction (pink).*

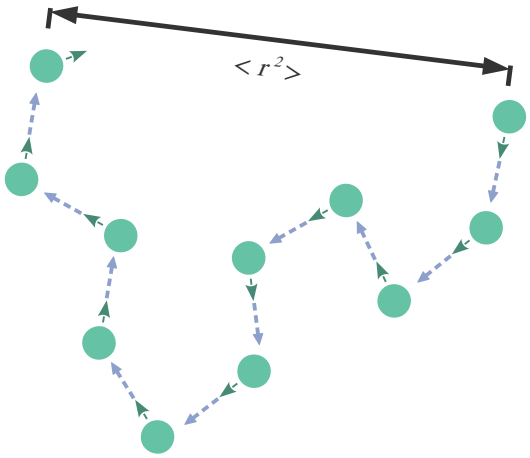


Figure 6.4 *Schematic demonstrating how a translational diffusion coefficient is determined for self-propelled particles which are not subject to direct brownian translational diffusion. The finite persistence time of their direction leads to diffusive behaviour at long times. The particle shown is performing run-and-tumble dynamics, however the same applies to a particle whose direction relaxes through continuous rotational diffusion.*

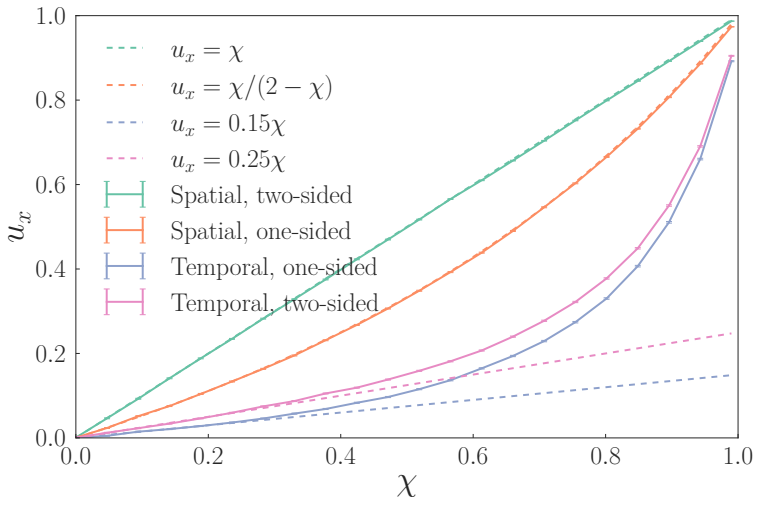


Figure 6.5 *Drift speed as a function of the chemotactic sensitivity for tumbling particles in one dimension. The spatial particles match the exact analytic predictions. Approximate fits to the temporal response in the linear, low-χ limit are indicated.*

speed of a diffusing particle doing spatial chemotaxis,

$$u_x(\chi) := \frac{v_x(\chi)}{v} = \frac{1-\sqrt{1-\chi^2}}{\chi}\,. \tag{6.14}$$

Figure 6.6 shows how the form of rotational noise affects the efficiency of migration for particles doing spatial chemotaxis. As can be seen, there is no discernible effect of switching between tumbling and slow diffusion. By approximate fitting, it seems that in the limit of low χ, a one-sided response reduces the migration efficiency by a factor of two, as in the one-dimensional case.

Figure 6.7 shows the equivalent data for particles doing temporal chemotaxis. In this case, the form of rotational noise becomes important. For $\chi \lesssim 0.5$ the response is similar to the spatial case, where the one-sided response is roughly half that of the two-sided response, and the form of rotational noise does not change the drift speed. Above this point, the tumbling particles have a higher drift speed than their diffusing counterparts. The size of this difference due to rotational noise increases as χ increases, such that for $\chi \gtrsim 0.8$, the form of noise is more important than the sided-ness of the response in predicting the migration efficiency.

A possible explanation for this difference can be reached by considering the reliability of the information gained by each type of particle at different points in

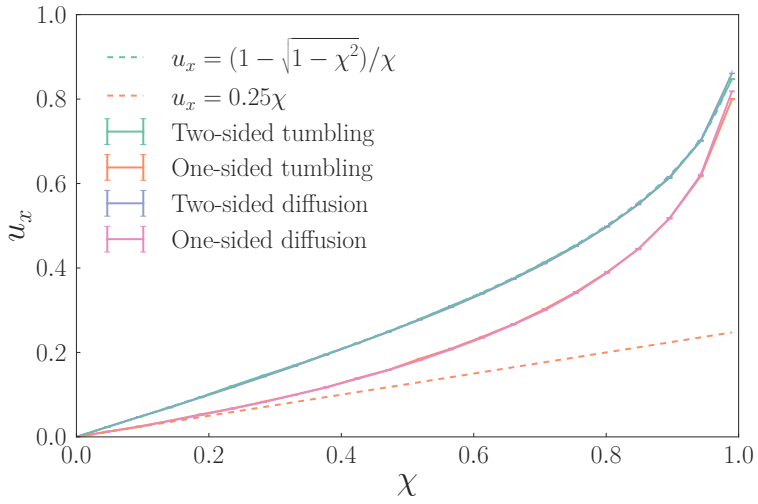


Figure 6.6 *Drift speed as a function of the chemotactic sensitivity for particles in two dimension doing spatial chemotaxis. The particles with a two-sided response match the exact analytic predictions. Approximate fits to the one-sided response in the linear, low-χ limit are indicated. The form of rotational noise has no effect on the efficiency of migration.*

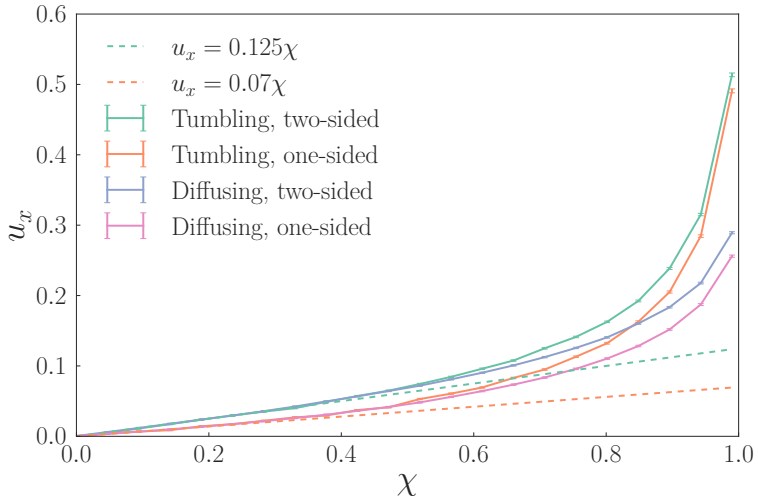


Figure 6.7 *Drift speed as a function of the chemotactic sensitivity for particles in two dimension doing temporal chemotaxis. Approximate fits to the tumbling responses in the linear, low-χ limit are indicated. At low χ, the effect of the sided-ness of the response dominates migration efficiency. For $\chi \gtrsim 0.8$, the form of rotational noise dominates.*

the past. The temporal method of doing chemotaxis depends on the measurement of chemical concentration over time being a good proxy for the spatial gradient. This is more accurate, the closer the particle's trajectory is to a straight line. Put another way, the particle's direction should remain similar over time, to estimate the gradient accurately. For a tumbling particle at some point in time, if we consider its direction one persistence-time ago, it is likely to have tumbled in this time, so its direction is likely to be entirely uncorrelated with its current direction, and any measured concentration changes before this tumble are of no value. However, in the recent past, the tumbling particle is likely not to have tumbled at all, and its path is therefore likely to be a perfectly straight line, and provides optimal gradient information. In the case of the slowly diffusing particle, its direction one persistence-time ago is more likely to have some correlation with its present direction than a tumbling particle, and it therefore provides some useful gradient information. Its path in the recent past is, however, less straight than that of a tumbling particle, and so provides less useful information. The tumbling method of rotational noise therefore provides more valuable information about the recent past, at the cost of less valuable information in the more distant past. When this is considered along with the kernel shown in Fig. 6.2, which assigns a large weight to the recent past, it seems reasonable that a tumbling particle might achieve a higher migration efficiency.

6.4 Model of a porous environment

We wish to investigate migration when the particle collide with obstacles. The environment we choose is one composed of a number spheres of constant radius. We vary the number of spheres and keep the radius constant, to vary the packing fraction. These environments are constructed using the Metropolis-Hastings algorithm outlined in Section 3.5. An example of such an environment is shown in Fig. 6.8. The radius we choose is $R = 20\,\mu\text{m}$, which is a sensible choice for environments such as dense sand [156–158], in a two-dimensional, square periodic system of size $L = 300\,\mu\text{m}$. We investigate packing fractions in the range ϕ =0 to 0.8, though it should be borne in mind that typically encountered packing fractions are in the range ϕ =0.2 to 0.5 [159].

We wish to investigate the effect of a swimmer's interaction with surfaces in a porous medium. Four distinct forms of interaction were implemented, as shown in Fig. 6.9. The aligning interaction is the relevant one for the flagellated bacteria

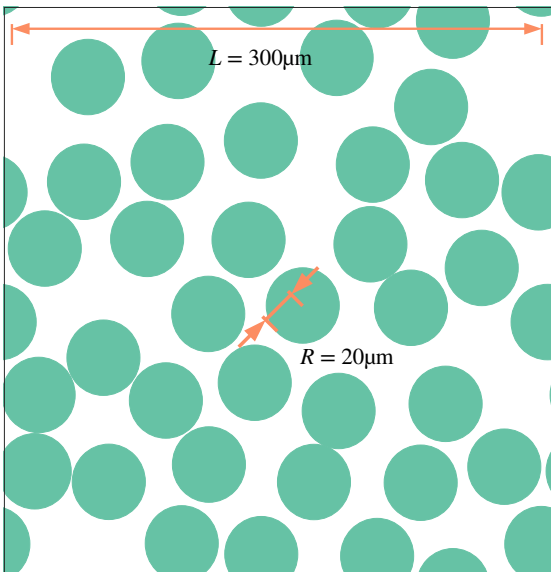


Figure 6.8 *An example of a typical periodic porous medium used in simulations. The environment shown has a packing fraction of $\phi = 0.5$. The circle sizes are as shown, and the system is square, with the period length shown along both axes.*

with which we are here concerned; alternatives are introduced to probe the properties of bacterial migration that depend on this interaction.

6.5 Diffusion in a porous environment

We first measure the ability of the swimmers to disperse in a porous medium with no chemical gradients. To investigate whether in such an environment, the dispersal of particles is diffusive, we measured the mean squared displacement of particles in a densely packed environment ($\phi = 0.8$), and measured the exponent, $\langle r^2 \rangle \propto t^{\Pi}$, as shown in Fig. 6.10. The largest deviation seen is for a reversing-type interaction, for which we see evidence of slightly sub-diffusive behaviour. As this is a small deviation even at the largest packing fractions we inspect, we assume behaviour is diffusive hereon.

Assuming diffusive behaviour, the diffusion coefficients in the porous media are shown in Fig. 6.11. An immediately counter-intuitive result that can be seen in Fig. 6.11a is that slowly diffusing swimmers that align with obstacles begin to increase their rate of permeation through a porous medium at large ϕ, while the motion of their tumbling counterparts continues to be inhibited. We were not able to investigate the cause of this discrepancy further, however we outline a possible

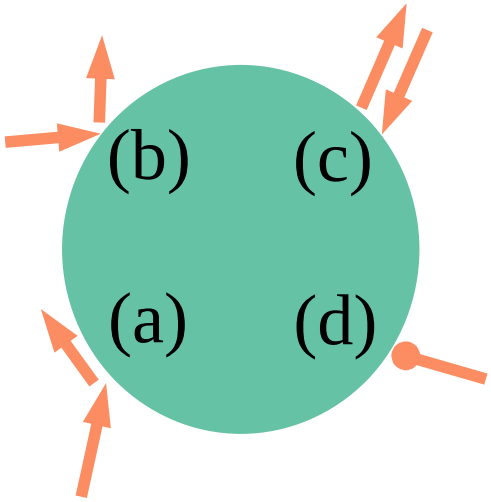


Figure 6.9 *Schematic of four possible interactions of a self-propelled particle with a surface. (a) Alignment: the particle aligns parallel with the surface and preserves its speed. (b) Reflection: the particle's velocity component perpendicular the surface reverses its sign. (c) Reversing: the swimmer's orientation is reversed on collision with a surface. (d) Stalling: The particle preserves its orientation but has zero speed.*

mechanism behind this difference. In an environment approaching the packing limit, the environment a particle must navigate becomes a network of narrow channels between obstacles, where the average distance to an obstacle is very small. For a particle whose direction changes continuously to reverse direction, its orientation must rotate at least 90 degrees over an extended period of time. During this time, a collision with an obstacle is likely to restore its direction to approximately the same direction as it began. This maintains a stronger degree of directional correlation over time, and therefore it diffuses faster. For tumbling particles, the necessary reversal of direction may occur instantaneously, so this 'guiding' effect of the environment is less pronounced. Even discarding this difference between modes of rotational noise, an aligning interaction allows for the most robust exploration of the environment.

It is interesting to note that the only other case where tumbling and diffusing particles are significantly distinguished in their behaviour is when they stall on encountering an obstacle (Fig. 6.11d). This is presumably because diffusing particles, whose direction continuously relaxes, are likely to leave the surface parallel with it, giving it a response more similar to the aligning interaction, while tumbling particles leave the surface through a tumbling event, and so have an isotropic distribution of leaving angles. The differentiation between the two

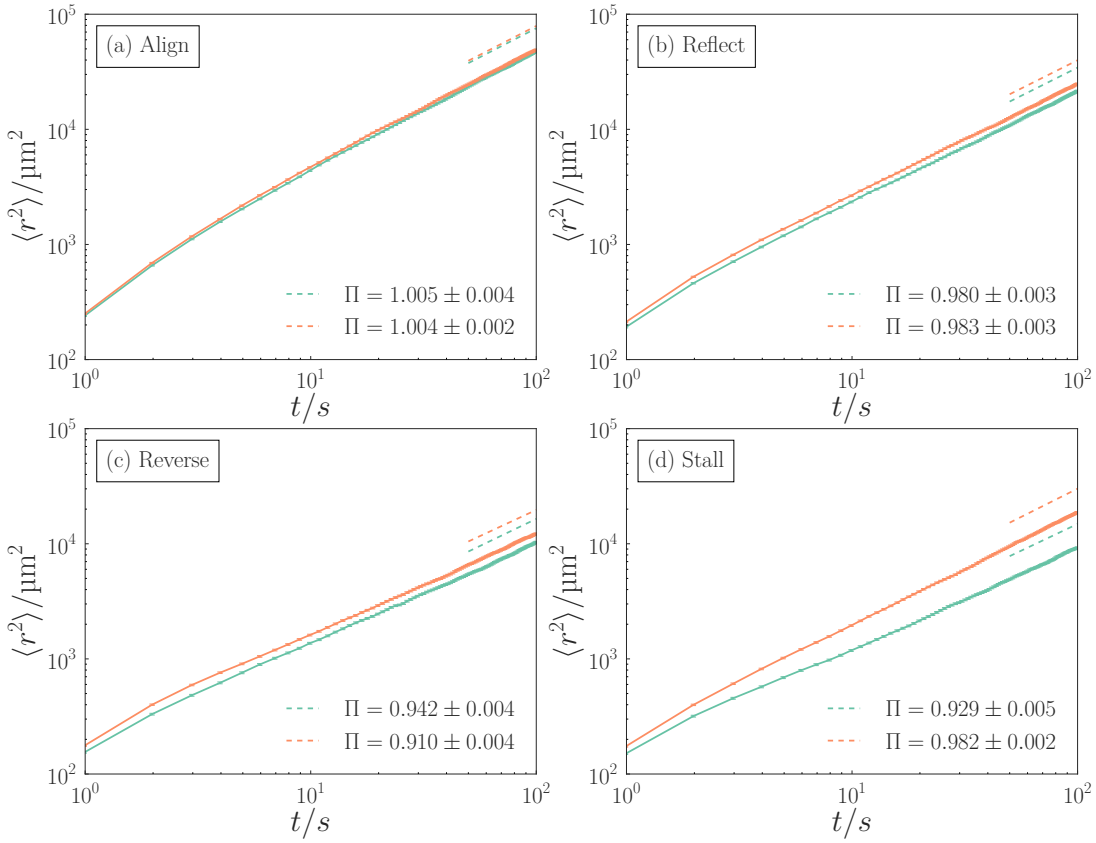


Figure 6.10 *The mean squared displacement of particles moving in a porous environment, with packing fraction $\phi = 0.8$. Particles' directions relax through either tumbling (green) or rotational diffusion (orange). We fit the curve to the equation $\langle r^2 \rangle = \mathrm{C}t^{\Pi}$ at long times to determine the diffusion characteristics of the particle, as it interacts with obstacles in the indicated manner. An aligning interaction leads to almost perfectly diffusive behaviour* $(\Pi = 1)$ *for both forms of rotational noise. The remainder show slightly sub-diffusive behaviour* $(\Pi < 1)$.

particle types is strong enough that the 'stalling' diffusing particles disperse faster than their 'reversing' counterparts (Fig. 6.11c), while the stalling tumbling particles disperse slower than their reversing counterparts.

We next investigate the effect of rotational noise on the ability of the particles to disperse through the environment, as in Fig. 6.3, in a porous medium. It might be expected that in such an environment, a small amount of rotational noise might enhance diffusivity, by aiding particles in navigating past obstacles. We have chosen a dense environment where $\phi = 0.5$ to make it easier to resolve this phenomenon, since it is more likely to occur when collisions are frequent. Figure 6.12a-b shows that for aligning and reflecting interactions, additional noise never aids particles in exploring the environment. The behaviour of tumbling and diffusing particles interacting with the environment in this way is identical. Up to a threshold, the degree of random rotational noise is irrelevant, as the particles' change in direction is dictated by collisions with the medium. Above this threshold, the diffusive behaviour becomes equivalent to the free-space case, except with a rate reduced by a constant factor.

For reversing and stalling particles, there is indeed a regime where additional rotational noise aids exploration. For reversing interactions, Fig. 6.11c implies that a porous environment does not distinguish between tumbling and diffusing particles, however from Fig. 6.12c it can be seen that in fact, for low levels of noise, tumbling particles are more sensitive to obstacles. Figure 6.12d indicates that for particles that stop at obstacles, the optimal level of noise for dispersal depends on the form of the noise, and that it is better to slowly diffuse the orientation.

6.6 Drift in a porous environment with a constant, linear gradient

From here onwards we focus on particles that align with obstacles, as it is the experimentally relevant case. We introduce a linear gradient of chemoattractant, increasing along the x-axis, and investigate how the migration of particles depends on their method of chemotaxis. As shown in Figs. 6.6 and 6.7, the same value of χ results in varying drift speeds in a uniform environment for different types of particle, so to make the models comparable, we choose an appropriate χ for each model to ensure an equal free-space drift speed (Table 6.1). The drift speed

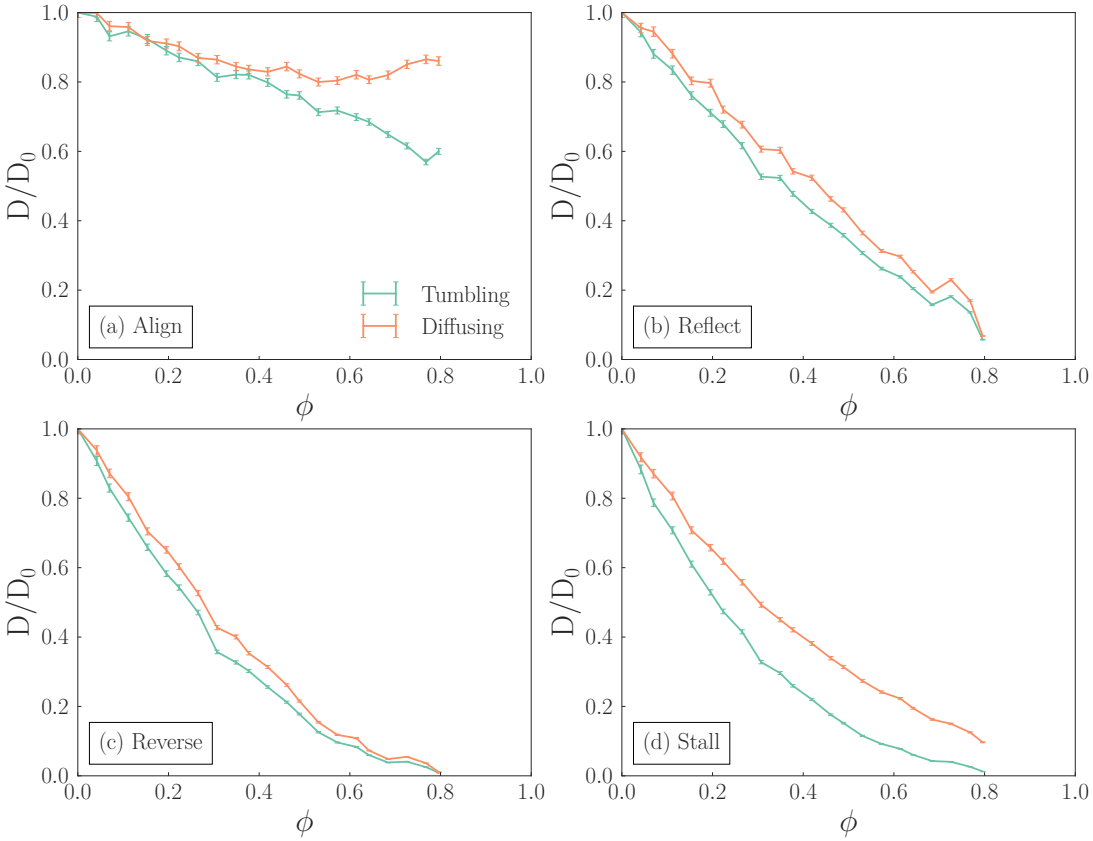


Figure 6.11 *The diffusivity,* D *of swimmers in a porous medium of variable packing fraction,* ϕ*. The diffusivity is normalised by the free-space diffusion,* $\mathrm{D}_0 = 400\,\mu\mathrm{m}^2/\mathrm{s}$*. (a) Swimmers which align parallel with surfaces. For* $\phi \gtrsim 0.4$*, the two models of rotational noise diverge, with the tumbling particles continuing to be inhibited in their dispersal, while the diffusive particles have their rate of dispersal enhanced in dense environments. (b) Swimmers which reflect their direction across the surface normal. Diffusive swimmers percolate through the medium slightly faster than tumbling swimmers, but the overall trend is a linear decrease. (c) Swimmers which reverse direction at a collision. Similarly to (b), Diffusive swimmers again percolate slightly faster than tumbling swimmers, but the inhibition of dispersal is stronger than for a reflective interaction. (d) Swimmers which stop swimming but do not change direction when coliiding with a surface. Compared with (c), tumbling particles are more strongly inhibited, while diffusing particles are less strongly inhibited.*

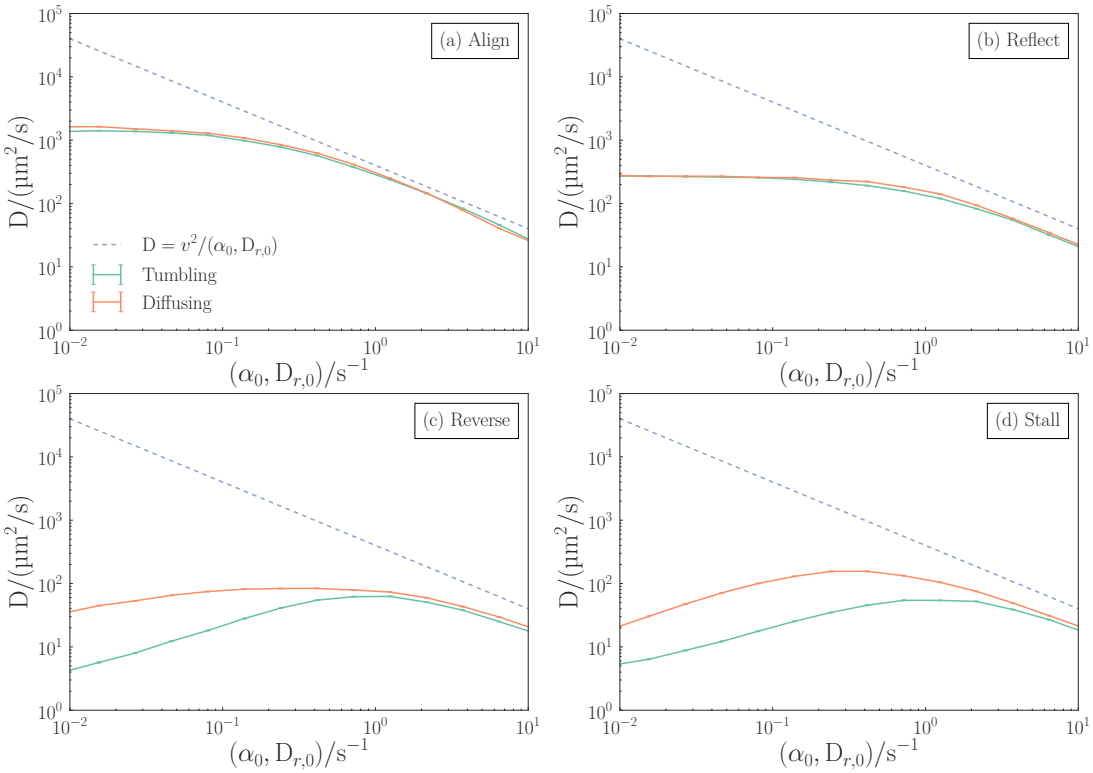


Figure 6.12 *Effective translational diffusion constant at long times for particles in a chemically uniform, porous medium, of packing fraction $\phi = 0.5$. The theoretical free-space diffusion constant is indicated (dashed line). Each plot shows a comparison between tumbling (green) and smoothly diffusing (orange) rotational noise, for a particular form of interaction with obstacles. (a) Alignment parallel with surfaces. There is no significant difference between forms of noise, and additional noise always inhibits the rate of dispersal. (b) Specular reflection, reversing the velocity component along the surface normal. Identical conclusion to aligning interactions are shown, but the overall rate of dispersal is slower. (c) Reversal of the velocity at collision. Tumbling particles are more strongly inhibited by the medium when it is the dominant source of velocity changes (at low α_0 or $\mathrm{D}_{r,0}$). Up to a threshold dependent on the form of noise, there is a region where additional rotational noise enhances dispersal through the environment (this threshold is higher for tumbling particles). (d) The particle stops on colliding with obstacles, while preserving its orientation. Similar conclusions are found as for reversing interactions. The tumbling particles respond identically whether they reverse or stall, while diffusing particles may diffuse faster or slower, depending on the strength of their rotational noise: reversal is preferred with low noise; stalling at intermediate noise.*

Estimation method	Noise	Response sidedness	χ
Spatial	Tumbling	2	0.385
Spatial	Diffusing	2	0.387
Spatial	Tumbling	1	0.548
Spatial	Diffusing	1	0.552
Temporal	Tumbling	2	0.855
Temporal	Diffusing	2	0.916
Temporal	Tumbling	1	0.887
Temporal	Diffusing	1	0.950

Table 6.1 *Required values of χ for each combination of model choices, in order to obtain a drift speed in a linear gradient of 20 % of the particle speed.*

we require is $u_{x,0} := v_{x,0}/v = 0.2$, where $v_{x,0}$ denotes the drift speed in free space. This is at the upper limit of the range of realistically achievable bacterial mobilities, which allows for meaningful conclusions, as well as good resolution of the drift speed in dense packings (which would be more difficult to measure with a smaller $u_{x,0}$).

The results for the two methods of estimating chemical gradients are shown in Fig. 6.13. Particles with a one- or two-sided response cope similarly with confinement, and the form of rotational noise also has little effect on the drop-off in migration efficiency. This degree of inhibition of particle migration may be quantified by the change in drift speed per change in packing fraction, normalised by the free-space drift speed to obtain a dimensionless number:

$$\beta := -\frac{\partial u_x/u_{x,0}}{\partial \phi} \,. \tag{6.15}$$

This quantity is a dimensionless measure, which represents the fraction by which the free-space drift speed is reduced, when the packing fraction of its environment is increased. For spatial chemotaxis, this gradient can be seen to remain approximately constant up to the largest ϕ measured, with values ranging from $\beta \simeq 0.45$ for two-sided diffusing particles, up to $\beta \simeq 0.65$ for one-sided tumbling particles. There must be a non-linear regime at higher ϕ where β increases, as $\beta < 1$ implies a non-zero drift speed in an environment that is entirely filled, which is not possible. For temporal chemotaxis, the response is less linear, but we estimate it for $\phi < 0.2$ as ranging from $\beta \simeq 1.15$ for two-sided diffusing particles, up to $\beta \simeq 1.95$ for one-sided tumbling particles. The larger values of β for particles doing temporal chemotaxis show that their ability to

migrate is restricted more by environmental obstacles than is the case for particles doing spatial chemotaxis, even when their drift speeds towards chemical gradients in free space are identical.

It is clear that temporal chemotaxis is more strongly inhibited in porous media, but no clear qualitative distinction can be drawn. What has been shown so far is that a porous medium decreases both the *dispersal* of bacteria, and also the *migration* of bacteria. It may be postulated that the response of these quantities are related, since they both originate from collisions with obstacles, and perhaps that they decrease at the same rate such that undirected random diffusion *around* an environment is always 'as fast' as directed migration *through* it.

To test this hypothesis, we calculate the ratio between the two transport properties, the diffusion constant, D, and the drift speed, u_x, normalised by their free-space values,

$$l_x(\phi) = \frac{\mathrm{D}(\phi)/\mathrm{D}_0}{u_x(\phi)/u_{x,0}} \,. \tag{6.16}$$

The fluctuation-dissipation theorem tells us that when a particular mechanism gives rise to a fluctuation in a particular variable, there is a connection to the dissipation of that same variable. We should expect from the theorem that drift and diffusion should be affected by collisions with obstacles at the same rate, and therefore that l_x should remain constant. Changes in this ratio therefore measure the degree of deviation from the fluctuation-dissipation theorem.

Figure 6.14 shows the estimate of this quantity over a range of packing fractions. For spatial chemotaxis, this ratio is indeed roughly constant over the range of ϕ investigated. The largest deviation from its free-space value is for one-sided, diffusing particles, where the increase is less than a factor of two at $\phi = 0.8$. For temporal chemotaxis, however, there is a significant increase. We perform a linear fit with an intercept at $l_x(\phi = 0) = 1$. For the one-sided tumbling case (the experimentally relevant one), the spatial fit gives a slope of $m = 0.43(3)$, while for the temporal case, $m = 2.01(9)$. It should be noted that in our analysis we have ignored the three data-points with largest ϕ for tumbling particles in Fig. 6.14b due to their difference from the pattern seen elsewhere. These values are not a sampling artifact as they are robust to increasing the length of the simulation and number of particles, but may be artifacts of the particular packing configuration used (however this has not been thoroughly investigated). Since

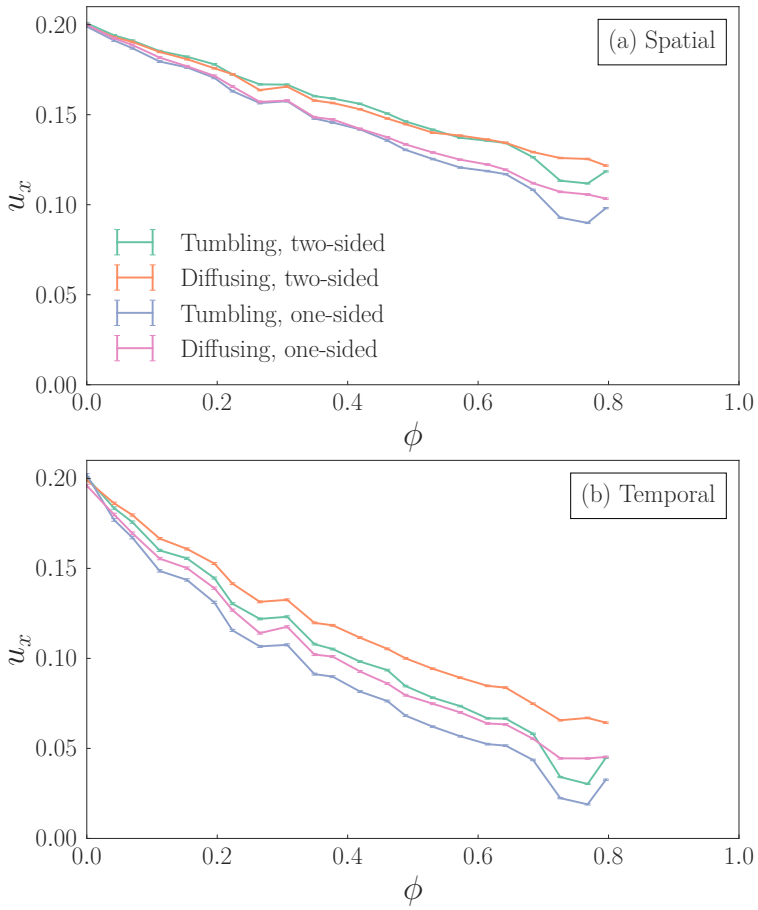


Figure 6.13 *The effect of a porous environment on the migration efficiency of chemotactic particles along the direction of increasing chemical concentration. The free-space drift speed is chosen to be* 20 % *of the particle speed in all cases. For the values of* χ *this implies, see Table 6.1. (a) Particles doing spatial chemotaxis. The form of rotational noise makes little difference to the inhibition of migration, while one-sided particles cope less well. (b) Particles doing temporal chemotaxis. The more negative slope in all cases compared with (a) shows that temporal chemotaxis is a less efficient method in environments where collisions are frequent. A two-sided response is less strongly inhibited, as its their spatial analogue, and slow diffusion performs better than tumbling, in contrast to the spatial case where these are not distinguishable.*

these high packing fractions are not physically realistic parameters, discarding this data is justified.

For self-propelled particles doing spatial chemotaxis, their approximately constant value of l_x indicates that their ability to navigate in a *particular* direction (drift) is made more difficult in proportion to the difficulty of moving in *any* direction (diffusion). The increase in l_x for self-propelled particles doing temporal chemotaxis demonstrates the presence of an additional mechanism which solely inhibits migration, but not dispersal, as this would cause $u_x(\phi)$ to decrease faster than $\mathrm{D}(\phi)$. We argue that this mechanism is the loss of useful gradient measurement information to the particle, because of a reduced correlation in its direction. Since temporal chemotaxis rests on maintaining a consistent direction long enough to measure concentration changes accurately, frequent collisions with the environment inhibit the particle's ability to measure chemical gradients. In summary, porous media prevent all self-propelled particles in getting anywhere; in addition, they prevent particles doing temporal chemotaxis from knowing where they *want* to go.

6.7 Conclusion

We have shown that in an environment that is free of obstacles, and for particles that are forced by their small scale to use temporal averaging to estimate chemical gradients, a method of rotational noise modulation that is implemented by discrete switching events (tumbling) is beneficial in aiding migration towards favourable regions. For particles in shallow chemical gradients, restricting the noise modulation to decrease the noise significantly reduces migration efficiency, however this difference becomes less important as the chemical gradients become more steep.

We have shown that in a porous medium, aligning parallel with obstacles is a good strategy for quickly exploring the environment, but the tumbling approach to modulating rotational noise leads to a reduced rate of exploration, regardless of what happens when particles meet an obstacle. We have also shown that if particles *do* align with obstacles, it is always in their interest to minimise their rate of rotational diffusion to disperse quickly. We have demonstrated a lack of monotonicity in the diffusion constant of swimmers with a slowly diffusing direction and which align with obstacles, as a function of the environment's porosity, which may deserve further study.

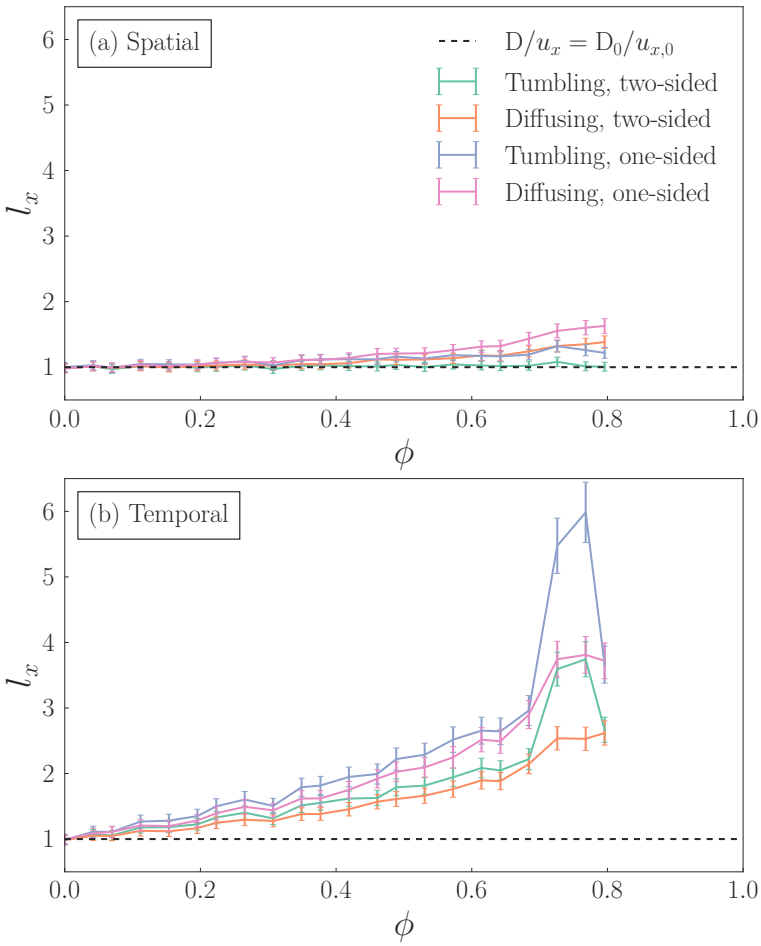


Figure 6.14 *The ratio of the diffusion constant of particles in a porous environment with no chemical gradients, to their drift speed when a linear gradient is present, as a function of the environment's packing fraction. The ratio is normalised by its value in free space, which is indicated for comparison (dashed line). (a) Particles doing spatial chemotaxis. The ratio remains close to its free-space value over the range of ϕ measured, with the largest deviation with one-sided diffusing particles at $\phi = 0.8$, where $l_x = 1.6(1)$. (b) Particles doing temporal chemotaxis. The same ratio in all cases rises much more rapidly over all ϕ. For comparison, one-sided diffusing particles at the same packing fraction has $l_x = 3.7(3)$, a rise roughly 6 times larger than in the spatial case.*

Finally, when migrating towards favourable chemicals in a porous medium, we have shown that collisions with the environment lead to significant deviations from the fluctuation-dissipation theorem at large packing fractions. The effective degree of noise induced by the environment is signicantly stronger for swimmers employing a temporal averaging method of chemotaxis, presumably due to the additional disruption of their estimate caused by collisions. This is exacerbated by a one-sided response, and additionally by a tumbling method of rotation. This implies that when considering the design of a self-propelled agent, to climb chemical gradients in a confined space, a small, flagellated prokaryote with the internal signalling mechanism outlined previously, is a bad example to follow.

Chapter 7

Conclusions

Self-propelled particles moving in free space can in many cases be mapped onto passive systems for which the important macroscopic properties can be inferred using standard techniques of statistical mechanics. Similarly, environments with spatial heterogeneities do not present a significant obstacle to analysis when the constituent particles under inspection are in thermal equilibrium. It is the combination of objects which exist out of equilibrium, and spatial heterogeneity, which leads to the novel phenomena we have described here. This thesis has demonstrated that the analysis of such systems is greatly aided by the detailed models that can be constructed through an agent-based approach, and makes a contribution to the understanding of how self-propelled particles respond to heterogeneous environments with complex boundaries. The models we have studied are simple enough to be relevant to non-equilibrium statistical physics, as we demonstrate examples of collective motion and an effective interaction between particles in the absence of any explicit attractive force. The models are nevertheless realistic enough to be relevant in a biological context, demonstrating the crucial impact of obstacles and collisions in determining the qualitative behaviour of bacteria.

We have processed and analysed experimental data studying the distribution of a smooth-swimming strain of *Escherichia. coli* suspended in the water phase of a water-in-oil emulsion, with spherical droplets of radii 5 µm to 40 µm at a range of bacterial densities. Our analysis showed two distinct regions of the droplet: an approximately 4 um shell close to the droplet interface where the density of bacteria was several times higher than the average, and the remainder of the droplet where

the density was uniform. After accounting for optical aberrations we fitted the data to an analytic model, from which was derived a parameter describing the degree to which a collision between bacteria disrupts their orientation. Agent based simulations were carried out to model the bacteria and the environment, in an attempt to reproduce the experimental data, purely from the microscopic parameters of the bacterial dynamics. The model reproduced the qualitative features of the experiments, and a fit to the analytic model provided an estimate of the degree of scattering experienced by colliding bacteria. This suggested that this scattering randomises the bacteria's orientation around 100 times faster than through Brownian rotational diffusion. We measured pair-wise correlations between bacterial positions in the droplet, which supported the prediction of strong scattering, that there are only weak pair-wise correlations.

It would be valuable in future work to extend the agent-based model to include flagellar components, and to apply the model in a cylindrical experimental setup to enable comparison with existing models and results in such a system. In a broader sense, our results motivate future experiments to image bacterial trajectories in detail, to resolve the exact form of their physical interactions. This may prove relevant to the early stages of biofilm formation, when motile organisms have not changed their expression profiles to a non-motile, communal state, and so dense, motile collections of organisms vie for space on surfaces.

We next investigated distributions of self-propelled particles at larger length scales and lower densities, where any interaction is mediated through a chemical concentration field. We ran agent-based simulations with point particles that contribute, through a linear source term, to a scalar field, and which modulate their rate of orientation randomisation in response to their orientation relative to the gradient of the same scalar field. We employed this model to investigate experimental observations of pattern formation in bacteria which secrete chemicals to which they exhibit a chemotactic response. We demonstrated a sharp transition to instability of a uniform distribution of particles, leading to smooth non-uniform distributions in one dimension, and singularities in two dimensions. To measure this transition we defined a robust measure of clustering using techniques from data analysis. We then introduced heterogeneities into the environment by adding obstacles to the system. We showed that this reduced the threshold for clustering, and broke symmetry in the location of cluster formation. We showed that environments presenting a large number of locations promoting density fluctuations can reduce the steady-state degree of clustering, by causing a large

number of local, stable density maxima to form which do not coalesce.

The complex interplay between surfaces and auto-chemotaxis is a promising avenue for research. The ready availability of micro-fabrication allows for tight feedback between simulations suggesting novel arrangements to control bacterial distributions, and experiments testing these hypotheses. There are practical outcomes in patterning surfaces to prevent or promote not only the presence of biofilms at all, but their particular locations in the environment, in human contexts such as medicine or waste treatment.

Finally, we investigated each microscopic aspect of the dynamics of chemotactic bacteria in turn, with an emphasis on their migration in porous media. We investigated the effect of implementing active rotational diffusion either discretely or continuously; the interaction with obstacles; whether the measurement of chemical gradients is sensed directly, or by averaging measurements over a finite time; and whether the response to this measurement is symmetric or not, with respect to the sign of the gradient. We showed that when migrating towards favourable regions using a time-averaged gradient measure, whether or not there is a response to both positive and negative gradients is the dominant determiner of efficiency, while in steep gradients, the method of rotational noise becomes the important feature. In porous media, we showed that discrete rotational noise (tumbling) is consistently an inferior method at diffusing through the environment, and migrating in a chemical gradient. We demonstrated the importance of velocity-aligning interactions for maintaining desirable transport properties in confined media, while providing intriguing evidence of a strong divergence between discrete and continuous rotational noise, when self-propelled particles align with surfaces. Our final result shows that the collisions caused by a confined environment lead to a greater reduction in migration efficiency when gradients are estimated as averages over time. We measure this by calculating deviations from the fluctuation-dissipation theorem, which shows that the rate of drift towards gradients is reduced by a much greater extent than is the rate of diffusion, in highly confined spaces, when time-averaging is done. This is proposed to be due to disruption of the particle's ability to estimate the true gradient, as its method assumes approximately straight paths.

The dependence of properties such as dispersal and migration efficiency on apparently minor details such as the method by which swimmers change direction, or their behaviour near surfaces, may have implications in the design of micro-swimmers. From a more theoretical perspective, it would be interesting to

investigate whether the complex effect of obstacles on bacteria doing chemotaxis by time-averaging, leads to a sharp transition in their migration efficiency, when chemical gradients are generated by the bacteria themselves. For example, in the common chemotactic wave assay, bacteria consume nearby nutrients, and in doing so generate gradients in these substances. It may be worthwhile to explore the effect of the density of the growth medium on chemotactic efficiency, as the work we have presented suggests non-linear responses may well be observed.

Appendix A

Derivation of expression for drift speed of active brownian particles doing spatial chemotaxis

Consider a single random walker moving on the unit circle, with a probability density of occupying a position θ, $P(\theta)$. The walker steps with a position-dependent step size, proportional to the local value of a field, $D(\theta)$. We assume the walker ergodically samples the circle, requiring $D(\theta) > 0 \forall \theta$. The dynamics of P may be expected to follow Fick's laws of diffusion, however generalising them to an inhomogeneous medium, where D varies in space, depends on the microscopic details of the diffusive process. The dependence of the change in probability density on the flux, $J(\theta, t)$, remains the same,

$$\frac{\partial P(\theta,t)}{\partial t} = -\frac{\partial J(\theta,t)}{\partial \theta}\,, \tag{A.1}$$

however the flux is no longer well predicted by its homogeneous form,

$$J(\theta,t) = -D\frac{\partial P(\theta,t)}{\partial \theta}\,. \tag{A.2}$$

An explanation of why the generalisation of Eq. (A.2) can depend on more than simply the form of D is given in [160]. There, the author presents the example of two boxes, each containing a dilute gas, and a dense wire mesh that inhibits the

motion of the gas particles. In the first box, the temperature is uniform, but the density of the wire mesh is not, so that the rate of diffusion of the gas particles varies within the box due to differing rates of collision with the wire. In the second box, the wire mesh is uniform, but the temperature is varied, so that the rate of diffusion varies due to differing particle speeds. The boxes are designed such that at all locations, their rates of diffusion are matched. Thermodynamics requires the equilibrium distribution of the first box to be uniform, since the temperature is constant. For the same reason, the second box must have density gradients. As such, knowledge of the diffusion constant alone is not enough to predict the dynamics of a diffusive system.

The first box has a diffusion rate which is limited by collisions; the second has one limited by speed. In our case of random walkers stepping on a circle, we argue that the correct interpretation is the speed-limited case, since our random walkers have no inertia, so the concept of a collision-induced direction change is not meaningful. In this case, the flux becomes,

$$J(\theta, t) = -\frac{\partial D(\theta) P(\theta, t)}{\partial \theta}\,, \tag{A.3}$$

which implies,

$$\frac{\partial P(\theta, t)}{\partial t} = \frac{\partial^2 D(\theta) P(\theta, t)}{\partial \theta^2}\,. \tag{A.4}$$

Taking this to steady-state,

$$P(\theta) = \frac{C + B\theta}{D(\theta)}\,, \tag{A.5}$$

where C and B are constants of integration. Since the only dependence on θ is implicitly through D, we can immediately conclude that $B = 0$.

As $P(\theta, t)$ is a probability density,

$$\int_{-\pi}^{\pi} P(\theta, t)\, \mathrm{d}\theta = 1\,, \tag{A.6}$$

so,

$$C^{-1} = \int_{-\pi}^{\pi} \frac{\mathrm{d}\theta}{D(\theta)}\,, \tag{A.7}$$

which allows the full probability density of the walker to be expressed in terms of the diffusion field:

$$P(\theta) = \left(\int_{-\pi}^{\pi} \frac{\mathrm{d}\theta}{D(\theta)} \right)^{-1} \frac{1}{D(\theta)}\,. \tag{A.8}$$

A.1 Drift speed for spatial chemotaxis

Using Eq. (A.8), the net drift in the direction $\theta = 0$ can be found,

$$\langle \cos(\theta) \rangle = \frac{\int_{-\pi}^{\pi} P(\theta) \cos(\theta)\, \mathrm{d}\theta}{\int_{-\pi}^{\pi} P(\theta)\, \mathrm{d}\theta} \tag{A.9}$$

$$\langle \cos(\theta) \rangle = \frac{\int_{-\pi}^{\pi} \dfrac{\cos(\theta)}{D(\theta)}\, \mathrm{d}\theta}{\int_{-\pi}^{\pi} \dfrac{\mathrm{d}\theta}{D(\theta)}} \tag{A.10}$$

The particular form of $D(\theta)$ for which we now solve is the form of spatial chemotaxis given in Eqs. (5.2) and (5.3), for a constant chemical field in the direction $\theta = 0$,

$$D(\theta) = D_0(1 - \chi \cos(\theta))\,, \tag{A.11}$$

which, when substituted into Eq. (A.10) gives,

$$\langle \cos(\theta) \rangle = \frac{\int_{-\pi}^{\pi} \dfrac{\cos(\theta)}{D_0(1 - \chi \cos(\theta))}\, \mathrm{d}\theta}{\int_{-\pi}^{\pi} \dfrac{\mathrm{d}\theta}{D_0(1 - \chi \cos(\theta))}}\,. \tag{A.12}$$

The absolute rate of diffusion does not affect the drift speed and cancels,

$$\langle \cos(\theta) \rangle = \frac{\int_{-\pi}^{\pi} \dfrac{\cos(\theta)}{1 - \chi \cos(\theta)} \, \mathrm{d}\theta}{\int_{-\pi}^{\pi} \dfrac{\mathrm{d}\theta}{1 - \chi \cos(\theta)}} \, . \tag{A.13}$$

For later convenience we express this as,

$$\langle \cos(\theta) \rangle = \frac{[I]_{-\pi}^{\pi}}{[I_N]_{-\pi}^{\pi}} \, . \tag{A.14}$$

Since it looks easier, we first evaluate the denominator of Eq. (A.14). By successive changes of variable, the result is found,

$$I_N = \int \frac{\mathrm{d}\theta}{1 - \chi \cos(\theta)} = -\frac{2}{\sqrt{\chi^2 - 1}} \tanh^{-1} \left(\frac{(\chi + 1) \tan\left(\theta/2\right)}{\sqrt{\chi^2 - 1}} \right) \, . \tag{A.15}$$

Evaluating this in the relevant limits,

$$[I_N]_{-\pi}^{\pi} = -\frac{2}{\sqrt{\chi^2 - 1}} \left[\tanh^{-1} \left(\frac{(\chi + 1) \tan\left(\theta/2\right)}{\sqrt{\chi^2 - 1}} \right) \right]_{-\pi}^{\pi} \, . \tag{A.16}$$

The fact that $\tanh^{-1}$ and tan are both odd functions allows this to be evaluated only at π,

$$[I_N]_{-\pi}^{\pi} = -\frac{4}{\sqrt{\chi^2 - 1}} \tanh^{-1} \left(\frac{(\chi + 1) \tan\left(\pi/2\right)}{\sqrt{\chi^2 - 1}} \right) \, . \tag{A.17}$$

Given that $\tan\left(\pi/2\right) = \infty$, and $\tanh^{-1}(\infty) = -i\pi/2$, then

$$[I_N]_{-\pi}^{\pi} = \frac{2\pi i}{\sqrt{\chi^2 - 1}} \, . \tag{A.18}$$

Since we require $\chi < 1$ for an ergodic system, we know that the lower term is

imaginary, and can factor out i, finally leaving,

$$[I_N]_{-\pi}^{\pi} = \frac{2\pi}{\sqrt{1-\chi^2}}\,. \tag{A.19}$$

We evaluate I using similar steps, as the relevant integral is of a similar form:

$$\int \mathrm{d}\theta \frac{\cos(\theta)}{1-\chi\cos(\theta)} = \frac{I_N - \theta}{\chi}\,, \tag{A.20}$$

leading to,

$$[I]_{-\pi}^{\pi} = \frac{[I_N]_{-\pi}^{\pi} - 2\pi}{\chi}\,, \tag{A.21}$$

which can be substituted back into Eq. (A.13) to give the final result,

$$\langle\cos(\theta)\rangle = \frac{1-\sqrt{1-\chi^2}}{\chi}\,. \tag{A.22}$$

Bibliography

[1] E. J. Marsden, C Valeriani, I Sullivan, M. E. Cates, and D Marenduzzo, Soft Matter **10**, 157 (2014).

[2] I. D. Vladescu, E. J. Marsden, J. Schwarz-Linek, V. A. Martinez, J Arlt, A. N. Morozov, D. Marenduzzo, M. E. Cates, and W. C. K. Poon, Physical Review Letters **113**, 268101 (2014).

[3] E. M. Purcell, Am. J. Phys **45**, 3 (1977).

[4] C. D. Silflow and P. A. Lefebvre, Plant physiology **127**, 1500 (2001).

[5] D. H. Lynn, *The ciliated protozoa: characterization, classification, and guide to the literature* (Springer Science \& Business Media, 2008).

[6] R. M. Macnab, Annual review of microbiology **57**, 77 (2003).

[7] F. Pacini, *Osservazioni microscopiche e deduzioni patologiche sul cholera asiatico* (tip. di F. Bencini, 1854).

[8] B. T. Rittenberg and S. C. Rittenberg, Archiv fi£¡r Mikrobiologie **42**, 138 (1962).

[9] A. Fàbrega and J. Vila, Clinical microbiology reviews **26**, 308 (2013).

[10] L. A. Pratt and R Kolter, Molecular Microbiology **30**, 285 (1998).

[11] G. Gauthier, M. Gauthier, and R. Christen, International Journal of Systematic Bacteriology **45**, 755 (1995).

[12] T Altindal, L Xie, and X. L. Wu, Biophysical Journal **100**, 32 (2011).

[13] P. S. Stewart, International Journal of Medical Microbiology **292**, 107 (2002).

[14] D. Gottlieb, The Journal of Antibiotics **29**, 987 (1976).

[15] J. W. Deacon and L. A. Berry, Pesticide Science **37**, 417 (1993).

[16] L. Hall-Stoodley, J. W. Costerton, and P. Stoodley, Nature Reviews Microbiology **2**, 95 (2004).

[17] K. K. Jefferson, FEMS Microbiology Letters **236**, 163 (2004).

[18] P Stoodley, I Dodds, J. D. Boyle, and H. M. Lappin-Scott, Journal of Applied Microbiology **85**, 19S (1998).

[19] J. W. Costerton, Z Lewandowski, D. E. Caldwell, D. R. Korber, and H. M. Lappin-Scott, Annual Review of Microbiology **49**, 711 (1995).

[20] R Di Leonardo, L Angelani, D DellâĂŹArciprete, G Ruocco, V Iebba, S Schippa, M. P. Conte, F Mecarini, F De Angelis, E Di Fabrizio, D Dell'arciprete, G Ruocco, V Iebba, S Schippa, M. P. Conte, F Mecarini, F De Angelis, and E Di Fabrizio, Proceedings of the National Academy of Sciences **107**, 9541 (2010).

[21] B. Joseph, S. Otta, I. Karunasagar, and I. Karunasagar, International Journal of Food Microbiology **64**, 367 (2001).

[22] D. G. Davies, M. R. Parsek, J. P. Pearson, B. H. Iglewski, J. W. Costerton, and E. P. Greenberg, Science **280**, 295 (1998).

[23] T.-F. C. Mah and G. A. O'Toole, Trends in Microbiology **9**, 34 (2001).

[24] D. G. Davies, A. M. Chakrabarty, and G. G. Geesey, Applied and Environmental Microbiology **59**, 1181 (1993).

[25] G. M. Teitzel and M. R. Parsek, Applied and Environmental Microbiology **69**, 2313 (2003).

[26] C. Prigent-Combaret, O. Vidal, C. Dorel, and P. Lejeune, J. Bacteriol. **181**, 5993 (1999).

[27] H.-C. Flemming, T. R. Neu, and D. J. Wozniak, Journal of bacteriology **189**, 7945 (2007).

[28] P. S. Stewart and J William Costerton, The Lancet **358**, 135 (2001).

[29] J. R. Bragg, R. C. Prince, E. J. Harner, and R. M. Atlas, Nature **368**, 413 (1994).

[30] R. Singh, D. Paul, and R. K. Jain, Trends in microbiology **14**, 389 (2006).

[31] M. M. Tunney, S. Patrick, M. D. Curran, G. Ramage, D. Hanna, J. R. Nixon, S. P. Gorman, R. I. Davis, and N. Anderson, J. Clin. Microbiol. **37**, 3281 (1999).

[32] A. V. Singh, V. Vyas, R. Patil, V. Sharma, P. E. Scopelliti, G. Bongiorno, A. Podestà, C. Lenardi, W. N. Gade, and P. Milani, PloS one **6**, e25029 (2011).

[33] P. D. Marsh, Caries research **38**, 204 (2004).

[34] P. I. Watnick and R. Kolter, Molecular Microbiology **34**, 586 (1999).

[35] K. J. Painter and J. A. Sherratt, Journal of Theoretical Biology **225**, 327 (2003).

[36] P. K. Singh, A. L. Schaefer, M. R. Parsek, T. O. Moninger, M. J. Welsh, and E. P. Greenberg, Nature **407**, 762 (2000).

[37] J. D. Shrout, D. L. Chopp, C. L. Just, M. Hentzer, M. Givskov, and M. R. Parsek, Molecular Microbiology **62**, 1264 (2006).

[38] L Rothschild, Nature **198**, 1221 (1963).

[39] A. P. Berke, L. Turner, H. C. Berg, and E. Lauga, Physical Review Letters **101**, 38102 (2008).

[40] G. Li and J. X. Tang, Physical Review Letters **103**, 78101 (2009).

[41] G. Li, J. Bensson, L. Nisimova, D. Munger, P. Mahautmr, J. X. Tang, M. R. Maxey, and Y. V. Brun, Physical Review E **84**, 41932 (2011).

[42] K. Drescher, J. Dunkel, L. H. Cisneros, S. Ganguly, and R. E. Goldstein, Proceedings of the National Academy of Sciences of the United States of America **108**, 10940 (2011).

[43] M. Molaei, M. Barry, R. Stocker, and J. Sheng, Physical Review Letters **113**, 068103 (2014).

[44] J. Prost, J.-F. Chauwin, L. Peliti, and A. Ajdari, Physical Review Letters **72**, 2652 (1994).

[45] M. E. Cates, Reports on Progress in Physics **75**, 42601 (2012).

[46] P. Galajda, J. E. Keymer, P. Chaikin, and R. H. Austin, Journal of Bacteriology **189**, 8704 (2007).

[47] J. Tailleur and M. E. Cates, Europhysics Letters **86**, 60002 (2009).

[48] A. Sokolov, I. S. Aranson, M. M. Apodaca, B. A. Grzybowski, and I. S. Aranson, Proceedings of the National Academy of Sciences **107**, 969 (2010).

[49] L Angelani and R Di Leonardo, Computer Physics Communications **182**, 1970 (2011).

[50] T. W. Engelmann, Pfl{ü}gers Archiv European Journal of Physiology **25**, 285 (1881).

[51] W. F. Pfeffer, *{Ü}ber chemotaktische bewegungen von Bacterien, Flagellaten und Volvocineen* (1888).

[52] H. C. Berg, *E. coli in Motion* (Springer Verlag, 2004).

[53] J. Adler, Science **153**, 708 (1966).

[54] J. Adler, Journal of General Microbiology **74**, 77 (1973).

[55] E. F. Keller and L. A. Segel, Journal of Theoretical Biology **30**, 235 (1971).

[56] F. A. C. C. Chalub, C. Schmeiser, B. Perthame, P. A. Markowich, B. Perthame, and C. Schmeiser, Monatshefte f{ü}r Mathematik **142**, 123 (2004).

[57] D Horstmann, Jahresber. Deutsch. Math.-Verein **105**, 103 (2003).

[58] M Holz and S. H. Chen, Biophysical Journal **26**, 243 (1979).

[59] J. D. Murray, *Mathematical Biology*, Vol. 2 (Springer, 2002).

[60] T. D. Brock, M. T. Madigan, J. M. Martinko, and J Parker, *Biology of microorganisms*, Vol. 112 (Prentice Hall London, UK, 1991).

[61] N. Mittal, E. O. Budrene, M. P. Brenner, and A. V. Oudenaarden, Proceedings of the National Academy of Sciences **100**, 13259 (2003).

[62] E. O. Budrene and H. C. Berg, Nature **349**, 630 (1991).

[63] E. O. Budrene and H. C. Berg, Nature **376**, 49 (1995).

[64] D. E. Woodward, R Tyson, M. R. Myerscough, and J. D. Murray, Biophysical Journal **68**, 2181 (1995).

[65] Y Blat and M Eisenbach, J. Bacteriol. **177**, 1683 (1995).

[66] S Park, P. M. Wolanin, E. A. Yuzbashyan, H Lin, N. C. Darnton, J. B. Stock, P. Silberzan, and R. H. Austin, Proceedings of the National Academy of Sciences **100**, 13910 (2003).

[67] J. Adler, Science **166**, 1588 (1969).

[68] S Park, P. M. Wolanin, E. A. Yuzbashyan, P. Silberzan, J. B. Stock, and R. H. Austin, Science **301**, 188 (2003).

[69] T. J. Newman and R. Grima, Physical Review E **70**, 051916 (2004).

[70] R. Grima, Physical Review Letters **95**, 128103 (2005).

[71] H. C. Berg and D. A. Brown, Nature **239**, 500 (1972).

[72] H. Szurmant and G. W. Ordal, Microbiology and molecular biology reviews : MMBR **68**, 301 (2004).

[73] I. B. Zhulin, A. F. Lois, and B. L. Taylor, FEBS Letters **367**, 180 (1995).

[74] H. Flores, E. Lobaton, S. Méndez-Diez, S. Tlupova, and R. Cortez, Bulletin of mathematical biology **67**, 137 (2005).

[75] A Einstein, Annalen der Physik **322**, 549 (1905).

[76] G. E. Uhlenbeck and L. S. Ornstein, Physical Review **36**, 823 (1930).

[77] H. Lamb, *Hydrodynamics 6th ed.*. (1932).

[78] J. Taktikos, "Modeling the random walk and chemotaxis of bacteria: Aspects of biofilm formation", PhD thesis (2013).

[79] V. F. Gaposhkin, Russian Mathematical Surveys **21**, 1 (1966).

[80] H. J. Butt and M Jaschke, Nanotechnology **6**, 1 (1995).

[81] G. K. Batchelor, *An Introduction to Fluid Dynamics* (Cambridge university press, 2000).

[82] B. P. Le Clair, A. E. Hamielec, and H. R. Pruppacher, Journal of the Atmospheric Sciences **27**, 308 (1970).

[83] M. Godin, A. K. Bryan, T. P. Burg, K. Babcock, and S. R. Manalis, Applied Physics Letters **91**, 123121 (2007).

[84] J. Happel and H. Brenner, *Low Reynolds number hydrodynamics: with special applications to particulate media*, Vol. 1 (Springer Science \& Business Media, 2012).

[85] S. Ramaswamy, The Mechanics and Statistics of Active Matter **1**, 323 (2010).

[86] M. C. Marchetti, J. F. Joanny, S. Ramaswamy, T. B. Liverpool, J. Prost, M. Rao, and R. A. Simha, Reviews of Modern Physics **85**, 1143 (2013).

[87] F. Schweitzer and J. D. Farmer, *Brownian agents and active particles: collective dynamics in the natural and social sciences* (Springer Science \& Business Media, 2007).

[88] J. Saragosti, P. Silberzan, and A. Buguin, PLOS One **7**, e35412 (2012).

[89] J. J. Uicker, G. R. Pennock, J. E. Shigley, and Others, *Theory of machines and mechanisms* (Oxford University Press Oxford, 2011).

[90] G. L. Hunter, K. V. Edmond, M. T. Elsesser, and E. R. Weeks, Opt. Express **19**, 17189 (2011).

[91] L. D. Landau and E. M. Lifshitz, Course of Theoretical Physics, 227 (1987).

[92] D. V. Nicolau, J. P. Armitage, and P. K. Maini, Computational Biology and Chemistry **33**, 269 (2009).

[93] J. Saragosti, V. Calvez, N Bournaveas, B. Perthame, A. Buguin, and P. Silberzan, Proceedings of the National Academy of Sciences **108**, 16235 (2011).

[94] G. B. Arfken, *Mathematical methods for physicists* (Academic press, 2013).

[95] P. S. Lovely and F. Dahlquist, Journal of Theoretical Biology **50**, 477 (1975).

[96] J. R. Howse, R. A. L. Jones, A. J. Ryan, T. Gough, R. Vafabakhsh, and R. Golestanian, Physical Review Letters **99**, 48102 (2007).

[97] A. J. Valocchi and M. Malmstead, Water Resources Research **28**, 1471 (1992).

[98] S. Descombes and M. Thalhammer, IMA Journal of Numerical Analysis **33**, 722 (2012).

[99] D Jacques, J SimÅŕnek, D Mallants, and M. T. van Genuchten, Journal of contaminant hydrology **88**, 197 (2006).

[100] C. Chiang, C. Dawson, and M. Wheeler, Transport in Porous Media **6** (1991).

[101] H. S. Rifai and P. B. Bedient, Water Resources Research **26**, 637 (1990).

[102] W. Kinzelbach, W. Schäfer, and J. Herzer, Water Resources Research **27**, 1123 (1991).

[103] W. F. Ames, *Numerical methods for partial differential equations* (Academic press, 2014).

[104] J. Crank and P. Nicolson, Advances in Computational Mathematics **6**, 207 (1996).

[105] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical recipes in C*, Vol. 2 (Cambridge university press Cambridge, 1996).

[106] D. D. L. Minh and D. L. P. Minh, Communications in Statistics - Simulation and Computation (2014).

[107] S. Chib and E. Greenberg, The American Statistician (2012).

[108] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, The Journal of Chemical Physics **21**, 1087 (1953).

[109] S. Labík and W. R. Smith, Molecular Simulation **12**, 23 (2006).

[110] W. W. Wood, The Journal of Chemical Physics **48**, 415 (1968).

[111] I. McDonald, Molecular Physics **23**, 41 (1972).

[112] M. Marchi and P. Procacci, The Journal of Chemical Physics **109**, 5194 (1998).

[113] S. Kirkpatrick, C. D. Gelatt, M. P. Vecchi, and Others, science **220**, 671 (1983).

[114] T Vicsek, A Czirók, E. Ben-Jacob, I. Cohen, O. Shochet, A Czir o k, E. Ben-Jacob, I. Cohen, and O. Shochet, Physical Review Letters **75**, 1226 (1995).

[115] G. Baglietto and E. V. Albano, Phys. Rev. E. Stat. Nonlin. Soft Matter Phys. **80**, 050103 (2009).

[116] H. Wioland, F. G. Woodhouse, J. Dunkel, J. O. Kessler, and R. E. Goldstein, Physical Review Letters **110**, 268102 (2013).

[117] K. M. Jones, H. Kobayashi, B. W. Davies, M. E. Taga, and G. C. Walker, Nat Rev Micro **5**, 619 (2007).

[118] B. D. L. Fitt, H. A. McCartney, and P. J. Walklate, Annual Review of Phytopathology **27**, 241 (2003).

[119] R. S. Mehrotra, *Plant Pathology* (McGraw-Hill Education (India) Pvt Limited, 2003).

[120] T.-C. Yang, Y.-W. Leu, H.-C. Chang-Chien, and R.-M. Hu, Journal of Bacteriology **191**, 2266 (2009).

[121] A. Stefan, G. Palazzo, A. Ceglie, E. Panzavolta, and A. Hochkoeppler, Biotechnology and Bioengineering **81**, 323 (2002).

[122] L. Baraban, F. Bertholle, M. L. M. Salverda, N. Bremond, P. Panizza, J. Baudry, J. A. G. M. de Visser, and J. Bibette, Lab Chip **11**, 4057 (2011).

[123] V. F. Aguiar, A. M. Donoghue, K Arsi, I Reyes-Herrera, J. H. Metcalf, F. S. de los Santos, P. J. Blore, and D. J. Donoghue, Foodborne Pathog. Dis. **10**, 435 (2013).

[124] L Chen and M Subirade, Biomaterials **27**, 4646 (2006).

[125] T. Ahmed and R. Stocker, Biophys. J. **95**, 4481 (2008).

[126] J. S. Parkinson and S. E. Houts, Journal of Bacteriology **151**, 106 (1982).

[127] R. Heim, A. B. Cubitt, and R. Y. Tsien, Nature **373**, 663 (1995).

[128] L. G. Wilson, V. A. Martinez, J Schwarz-Linek, J Tailleur, G Bryant, P. N. Pusey, and W. C. K. Poon, Physical Review Letters **106**, 18101 (2011).

[129] V. A. Martinez, R. Besseling, O. A. Croze, J. Tailleur, M. Reufer, J. Schwarz-Linek, L. G. Wilson, M. A. Bees, and W. C. K. Poon, Biophysical Journal **103**, 1637 (2012).

[130] H. E. Kubitschek, Journal of Bacteriology **172**, 94 (1990).

[131] V. J. Lumelsky, Information Processing Letters **21**, 55 (1985).

[132] A. Sokolov, I. S. Aranson, J. O. Kessler, R. E. Goldstein, A. Sokolov, J. O. Kessler, R. E. Goldstein, I. S. Aranson, J. O. Kessler, and R. E. Goldstein, Physical Review Letters **98**, 158102 (2007).

[133] H. C. Berg, *Random walks in biology* (Princeton Univ Pr, 1993).

[134] R. M. Neal, The Annals of Statistics **31**, 705 (2003).

[135] N. F. Carnahan and K. E. Starling, The Journal of Chemical Physics **51**, 635 (1969).

[136] Q. Liao, G. Subramanian, M. P. DeLisa, D. L. Koch, and M. Wu, Physics of Fluids (1994-present) **19**, 61701 (2007).

[137] K. Gade, Journal of Navigation **63**, 395 (2010).

[138] G. T. McCaw, Survey Review **1**, 259 (1932).

[139] J. S. Parkinson, S. R. Parker, P. B. Talbert, and S. E. Houts, Journal of Bacteriology **155**, 265 (1983).

[140] N Vladimirov, V. Sourjik, and N Vladimirov, Biological Chemistry **390**, 1097 (2009).

[141] M. W. Germann, T. Turner, and S. A. Allison, The Journal of Physical Chemistry A **111**, 1452 (2007).

[142] D. Comaniciu, P. Meer, and S. Member, Pattern Analysis and Machine Intelligence, IEEE Transactions on **24**, 603 (2002).

[143] S. E. Prochnik, J. Umen, A. M. Nedelcu, A. Hallmann, S. M. Miller, I. Nishii, P. Ferris, A. Kuo, T. Mitros, L. K. Fritz-Laylin, U. Hellsten, J. Chapman, O. Simakov, S. A. Rensing, A. Terry, J. Pangilinan, V. Kapitonov, J. Jurka, A. Salamov, H. Shapiro, J. Schmutz, J. Grimwood, E. Lindquist, S. Lucas, I. V. Grigoriev, R. Schmitt, D. Kirk, and D. S. Rokhsar, Science **329**, 223 (2010).

[144] M. B. Wan and C. J. O. Reichhardt, Physical Review Letters, 4 (2008).

[145] A. Ringrose-Voase and W. Sanidad, Geoderma **71**, 245 (1996).

[146] W Brutsaert, Soil Science **101** (1966).

[147] S. Gupta Sood, FEMS microbiology ecology **45**, 219 (2003).

[148] Y. Bashan, G. Holguin, and L. E. De-Bashan, Canadian journal of microbiology **50**, 521 (2004).

[149] R. Thar and M. Kuhl, Proceedings of the National Academy of Sciences of the United States of America **100**, 5748 (2003).

[150] V. Sourjik, Trends in microbiology **12**, 569 (2004).

[151] P. Engström and G. L. Hazelbauer, Cell **20**, 165 (1980).

[152] J. A. Gegner, D. R. Graham, A. F. Roth, and F. W. Dahlquist, Cell **70**, 975 (1992).

[153] J. V. Olsen, B. Blagoev, F. Gnad, B. Macek, C. Kumar, P. Mortensen, and M. Mann, Cell **127**, 635 (2006).

[154] S. J. Kleene, A. C. Hobson, and J. Adler, Proceedings of the National Academy of Sciences **76**, 6309 (1979).

[155] D. A. Clark and L. C. Grant, Proceedings of the National Academy of Sciences **102**, 9150 (2005).

[156] J. Barton and R. Ford, Appl. Envir. Microbiol. **61**, 3329 (1995).

[157] P. Vandevivere and P. Baveye, Soil Science Society of America Journal **56**, 1 (1992).

[158] J. W. Barton and R. M. Ford, Biotechnology and bioengineering **53**, 487 (1997).

[159] P. K. W. D. C. Beard, AAPG Bulletin **57**, 349 (1973).

[160] M. J. Schnitzer, Physical Review E **48**, 2553 (1993).